PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/86, 15/12, A61K 48/00

A2

(11) International Publication Number: WO 94/12649

(43) International Publication Date: 9 June 1994 (09.06.94)

(21) International Application Number: PCT/US93/11667

(22) International Filing Date: 2 December 1993 (02.12.93)

(30) Priority Data:

07/985,478	3 December 1992 (03.12.92)	US
08/130,682	1 October 1993 (01.10.93)	US
08/136,742	13 October 1993 (13.10.93)	US

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(71) Applicant: GENZYME CORPORATION [US/US]; One Kendall Square, Cambridge, MA 02139 (US).

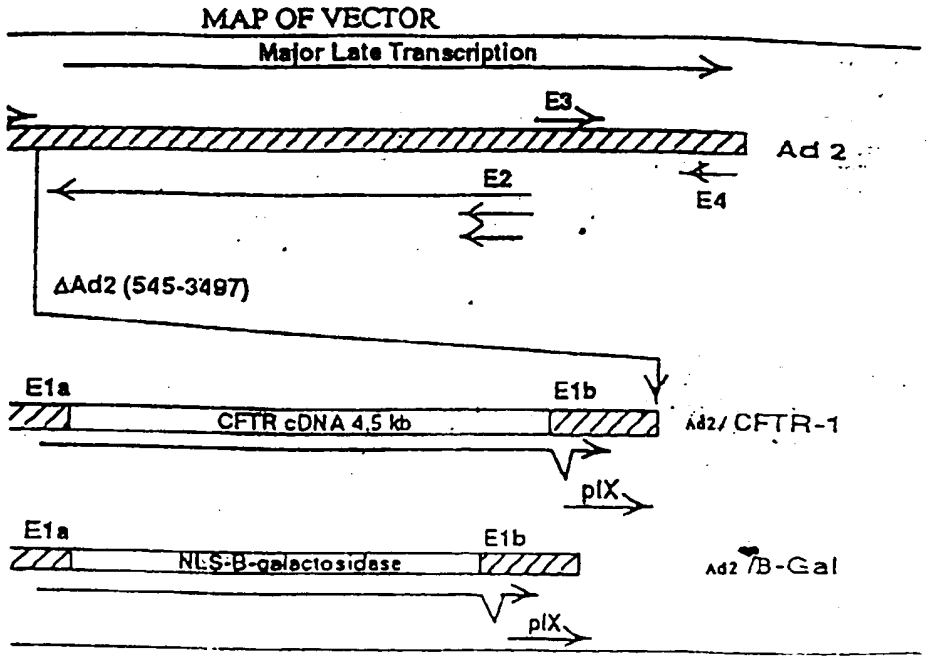
(72) Inventors: GREGORY, Richard, J.; 4789 Gateshead Road, Carlsbad, CA 92008 (US). ARMENTANO, Donna; 33 Carver Road, Watertown, MA 02172 (US). COUTURE, Larry, A.; 67 Circle Drive, Framingham, MA 01701 (US). SMITH, Alan, E.; 88 Cleveland Road, Wellesley, MA 02181 (US).

(74) Agents: HANLEY, Elizabeth, A. et al.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US).

(54) Title: GENE THERAPY FOR CYSTIC FIBROSIS

(57) Abstract

Gene Therapy vectors, which are especially useful for cystic fibrosis, and methods for using the vectors are disclosed. In preferred embodiments, the vectors are adenovirus-based. Advantages of adenovirus-based vectors for gene therapy are that they appear to be relatively safe and can be manipulated to encode the desired gene product and at the same time are inactivated in terms of their ability to replicate in a normal lytic viral life cycle. Additionally, adenovirus has a natural tropism for airway epithelia. Therefore, adenovirus-based vectors are particularly preferred for respiratory gene therapy applications such as gene therapy for cystic fibrosis. In one embodiment, the adenovirus-based gene therapy vector comprises an adenovirus 2 serotype genome in which the E1a and E1b regions of the genome, which are involved in early stages of viral replication have been deleted and replaced by genetic material of interest (e.g., DNA encoding the cystic fibrosis transmembrane regulator protein). In another embodiment, the adenovirus-based therapy vector is a pseudo-adenovirus (PAV). PAVs contain no potentially harmful viral genes, have a theoretical capacity for foreign material of nearly 36 kb, may be produced in reasonably high titers and maintain the tropism of the parent adenovirus for dividing and non-dividing human target cell types.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

GENE THERAPY FOR CYSTIC FIBROSIS

Related Applications

This application is a continuation-in-part application of United States Serial Number 08/130,682, filed on October 1, 1993 which is a continuation-in-part application of United States Serial Number 07/985,478, filed on December 2, 1992, which is a continuation-in-part application of United States Serial Number 07/613,592, filed on November 15, 1990, which is in turn a continuation-in-part application of United States Serial Number 07/589,295, filed on September 27, 1990, which is itself a continuation-in-part application of United States Serial Number 07/488,307, filed on March 5, 1990. The contents of all of the above co-pending patent applications are incorporated herein by reference. Definitions of language or terms not provided in the present application are the same as those set forth in the copending applications. Any reagents or materials used in the examples of the present application whose source is not expressly identified also is the same as those described in the copending application, e.g., ΔF508 CFTR gene and CFTR antibodies.

Background of the Invention

Cystic Fibrosis (CF) is the most common fatal genetic disease in humans (Boat, T.F. et al. in The Metabolic Basis of Inherited Diseases (Scriver, C.R. et al. eds., McGraw-Hill, New York (1989)). Approximately one in every 2,500 infants in the United States is born with the disease. At the present time, there are approximately 30,000 CF patients in the United States. Despite current standard therapy, the median age of survival is only 26 years. Disease of the pulmonary airways is the major cause of morbidity and is responsible for 95% of the mortality. The first manifestation of lung disease is often a cough, followed by progressive dyspnea. Tenacious sputum becomes purulent because of colonization of Staphylococcus and then with Pseudomonas. Chronic bronchitis and bronchiectasis can be partially treated with current therapy, but the course is punctuated by increasingly frequent exacerbations of the pulmonary disease. As the disease progresses, the patient's activity is progressively limited. End-stage lung disease is heralded by increasing hypoxemia, pulmonary hypertension, and cor pulmonale.

The upper airways of the nose and sinuses are also involved in CF. Most patients with CF develop chronic sinusitis. Nasal polyps occur in 15-20% of patients and are common by the second decade of life. Gastrointestinal problems are also frequent in CF; infants may suffer meconium ileus. Exocrine pancreatic insufficiency, which produces symptoms of malabsorption, is present in the large majority of patients with CF. Males are almost uniformly infertile and fertility is decreased in females.

Based on both genetic and molecular analyses, a gene associated with CF was isolated as part of 21 individual cDNA clones and its protein product predicted (Kerem, B.S. et al. (1989) *Science* 245:1073-1080; Riordan, J.R. et al. (1989) *Science* 245:1066-1073;

Rommens, J.M. et al. (1989) *Science* 245:1059-1065)). United States Serial Number 07/488,307 describes the construction of the gene into a continuous strand, expression of the gene as a functional protein and confirmation that mutations of the gene are responsible for CF. (*See also* Gregory, R.J. et al. (1990) *Nature* 347:382-386; Rich, D.P. et al. (1990) *Nature* 347:358-362). The co-pending patent application also discloses experiments which show that proteins expressed from wild type but not a mutant version of the cDNA complemented the defect in the cAMP regulated chloride channel shown previously to be characteristic of CF.

The protein product of the CF associated gene is called the cystic fibrosis transmembrane conductance regulator (CFTR) (Riordan, J.R. et al. (1989) *Science* 245:1066-1073). CFTR is a protein of approximately 1480 amino acids made up of two repeated elements, each comprising six transmembrane segments and a nucleotide binding domain. The two repeats are separated by a large, polar, so-called R-domain containing multiple potential phosphorylation sites. Based on its predicted domain structure, CFTR is a member of a class of related proteins which includes the multi-drug resistance (MDR) or P-glycoprotein, bovine adenyl cyclase, the yeast STE6 protein as well as several bacterial amino acid transport proteins (Riordan, J.R. et al. (1989) *Science* 245:1066-1073; Hyde, S.C. et al. (1990) *Nature* 346:362-365). Proteins in this group, characteristically, are involved in pumping molecules into or out of cells.

CFTR has been postulated to regulate the outward flow of anions from epithelial cells in response to phosphorylation by cyclic AMP-dependent protein kinase or protein kinase C (Riordan, J.R. et al. (1989) *Science* 245:1066-1073; Welsh, 1986; Frizzell, R.A. et al. (1986) *Science* 233:558-560; Welsh, M.J. and Liedtke, C.M. (1986) *Nature* 322:467; Li, M. et al. (1988) *Nature* 331:358-360; Huang, T-C. et al. (1989) *Science* 244:1351-1353).

Sequence analysis of the CFTR gene of CF chromosomes has revealed a variety of mutations (Cutting, G.R. et al. (1990) *Nature* 346:366-369; Dean, M. et al. (1990) *Cell* 61:863-870; and Kerem, B-S. et al. (1989) *Science* 245:1073-1080; Kerem, B-S. et al. (1990) *Proc. Natl. Acad. Sci.* USA 87:8447-8451). Population studies have indicated that the most common CF mutation, a deletion of the 3 nucleotides that encode phenylalanine at position 508 of the CFTR amino acid sequence (ΔF508), is associated with approximately 70% of the cases of cystic fibrosis. This mutation results in the failure of an epithelial cell chloride channel to respond to cAMP (Frizzell R.A. et al. (1986) *Science* 233:558-560; Welsh, M.J. (1986) *Science* 232:1648-1650.; Li, M. et al. (1988) *Nature* 331:358-360; Quinton, P.M. (1989) *Clin. Chem.* 35:726-730). In airway cells, this leads to an imbalance in ion and fluid transport. It is widely believed that this causes abnormal mucus secretion, and ultimately results in pulmonary infection and epithelial cell damage.

Studies on the biosynthesis (Cheng, S.H. et al. (1990) *Cell* 63:827-834; Gregory, R.J. et al. (1991) *Mol. Cell Biol.* 11:3886-3893) and localization (Denning, G.M. et al. (1992) *J. Cell Biol.* 118:551-559) of CFTR ΔF508, as well as other CFTR mutants, indicate that many CFTR mutant proteins are not processed correctly and, as a result, are not delivered to the

plasma membrane (Gregory, R.J. et al. (1991) *Mol. Cell Biol.* 11:3886-3893). These conclusions are consistent with earlier functional studies which failed to detect cAMP-stimulated Cl^- channels in cells expressing CFTR ΔF508 (Rich, D.P. et al. (1990) *Nature* 347:358-363; Anderson, M.P. et al. (1991) *Science* 251:679-682).

To date, the primary objectives of treatment for CF have been to control infection, promote mucus clearance, and improve nutrition (Boat, T.F. et al. in The Metabolic Basis of Inherited Diseases (Scriver, C.R. et al. eds., McGraw-Hill, New York (1989)). Intensive antibiotic use and a program of postural drainage with chest percussion are the mainstays of therapy. However, as the disease progresses, frequent hospitalizations are required. Nutritional regimens include pancreatic enzymes and fat-soluble vitamins. Bronchodilators are used at times. Corticosteroids have been used to reduce inflammation, but they may produce significant adverse effects and their benefits are not certain. In extreme cases, lung transplantation is sometimes attempted (Marshall, S. et al. (1990) *Chest* 98:1488).

Most efforts to develop new therapies for CF have focused on the pulmonary complications. Because CF mucus consists of a high concentration of DNA, derived from lysed neutrophils, one approach has been to develop recombinant human DNase (Shak, S. et al. (1990) *Proc. Natl. Sci. Acad USA* 87:9188). Preliminary reports suggest that aerosolized enzyme may be effective in reducing the viscosity of mucus. This could be helpful in clearing the airways of obstruction and perhaps in reducing infections. In an attempt to limit damage caused by an excess of neutrophil derived elastase, protease inhibitors have been tested. For example, alpha-1-antitrypsin purified from human plasma has been aerosolized to deliver enzyme activity to lungs of CF patients (McElvaney, N. et al. (1991) *The Lancet* 337:392). Another approach would be the use of agents to inhibit the action of oxidants derived from neutrophils. Although biochemical parameters have been successfully measured, the long term beneficial effects of these treatments have not been established.

Using a different rationale, other investigators have attempted to use pharmacological agents to reverse the abnormally decreased chloride secretion and increased sodium absorption in CF airways. Defective electrolyte transport by airway epithelia is thought to alter the composition of the respiratory secretions and mucus (Boat, T.F. et al. in The Metabolic Basis of Inherited Diseases (Scriver, C.R. et al. eds., McGraw-Hill, New York (1989); Quinton, P.M. (1990) *FASEB J.* 4:2709-2717). Hence, pharmacological treatments aimed at correcting the abnormalities in electrolyte transport could be beneficial. Trials are in progress with aerosolized versions of the drug amiloride; amiloride is a diuretic that inhibits sodium channels, thereby inhibiting sodium absorption. Initial results indicate that the drug is safe and suggest a slight change in the rate of disease progression, as measured by lung function tests (Knowles, M. et al. (1990) *N. Eng. J. Med.* 322: 1189-1194; App, E.(1990) *Am. Rev. Respir. Dis.* 141:605). Nucleotides, such as ATP or UTP, stimulate purinergic receptors in the airway epithelium. As a result, they open a class of chloride channel that is different from CFTR chloride channels. *In vitro* studies indicate that ATP and UTP can stimulate

chloride secretion (Knowles, M. et al. (1991) *N. Eng. J. Med.* 325:533). Preliminary trials to test the ability of nucleotides to stimulate secretion *in vivo*, and thereby correct the electrolyte transport abnormalities are underway.

Despite progress in therapy, cystic fibrosis remains a lethal disease, and no current therapy treats the basic defect. However, two general approaches may prove feasible. These are: 1) protein replacement therapy to deliver the wild type protein to patients to augment their defective protein, and; 2) gene replacement therapy to deliver wild type copies of the CF associated gene. Since the most life threatening manifestations of CF involve pulmonary complications, epithelial cells of the upper airways are appropriate target cells for therapy.

The feasibility of gene therapy has been established by introducing a wild type cDNA into epithelial cells from a CF patient and demonstrating complementation of the hallmark defect in chloride ion transport (Rich, D.P. et al. (1990) *Nature* 347:358-363). This initial work involved cells in tissue culture, however, subsequent work has shown that to deliver the gene to the airways of whole animals, defective adenoviruses may be useful (Rosenfeld, (1992) *Cell* 68:143-155). However, the safety and effectiveness of using defective adenoviruses remain to be demonstrated.

Summary of the Invention

In general, the instant invention relates to vectors for transferring selected genetic material of interest (e.g., DNA or RNA) to cells *in vivo*. In preferred embodiments, the vectors are adenovirus-based. Advantages of adenovirus-based vectors for gene therapy are that they appear to be relatively safe and can be manipulated to encode the desired gene product and at the same time are inactivated in terms of their ability to replicate in a normal lytic viral life cycle. Additionally, adenovirus has a natural tropism for airway epithelia. Therefore, adenovirus-based vectors are particularly preferred for respiratory gene therapy applications such as gene therapy for cystic fibrosis.

In one embodiment, the adenovirus-based gene therapy vector comprises an adenovirus 2 serotype genome in which the Ela and Elb regions of the genome, which are involved in early stages of viral replication have been deleted and replaced by genetic material of interest (e.g., DNA encoding the cystic fibrosis transmembrane regulator protein).

In another embodiment, the adenovirus-based therapy vector is a pseudo-adenovirus (PAV). PAVs contain no potentially harmful viral genes, have a theoretical capacity for foreign material of nearly 36 kb, may be produced in reasonably high titers and maintain the tropism of the parent adenovirus for dividing and non-dividing human target cell types. PAVs comprise adenovirus inverted terminal repeats and the minimal sequences of a wild-type adenovirus type 2 genome necessary for efficient replication and packaging by a helper virus and genetic material of interest. In a preferred embodiment, the PAV contains adenovirus 2 sequences.

In a further embodiment, the adenovirus-based gene therapy vector contains the open reading frame 6 (ORF6) of adenoviral early region 4 (E4) from the E4 promoter and is deleted for all other E4 open reading frames. Optionally, this vector can include deletions in the E1 and/or E3 regions. Alternatively, the adenovirus-based gene therapy vector contains the open reading frame 3 (ORF3) of adenoviral E4 from the E4 promoter and is deleted for all other E4 open reading frames. Again, optionally, this vector can include deletions in the E1 and/or E3 regions. The deletion of non-essential open reading frames of E4 increases the cloning capacity by approximately 2 kb without significantly reducing the viability of the virus in cell culture. In combination with deletions in the E1 and/or E3 regions of adenovirus vectors, the theoretical insert capacity of the resultant vectors is increased to 8-9 kb.

The invention also relates to methods of gene therapy using the disclosed vectors and genetically engineered cells produced by the method.

Brief Description of the Tables and Drawings

Further understanding of the invention may be had by reference to the tables and figures wherein:

Table I shows CFTR mutants wherein the known association with CF (Y, yes or N, no), exon localization, domain location and presence (+) or absence (-) of bands A, B, and C of mutant CFTR species is shown. TM6, indicates transmembrane domain 6; NBD nucleotide binding domain; ECD, extracellular domain and Term, termination at 21 codons past residue 1337;

Table II shows the nucleotide sequence of Ad2/CFTR-1;

Table III depicts a nucleotide analysis of Ad2-ORF6/PGK-CFTR;

The convention for naming mutants is first the amino acid normally found at the particular residue, the residue number (Riordan, T.R. et al. (1989) *Science* 245:1066-1073). and the amino acid to which the residue was converted. The single letter amino acid code is used: D, aspartic acid; F, phenylalanine; G, glycine; I, isoleucine; K, lysine; M, methionine; N, asparagine; Q, glutamine; R, arginine; S, serine; W, tryptophan. Thus G551D is a mutant in which glycine 551 is converted to aspartic acid;

Figure 1 shows alignment of CFTR partial cDNA clones used in construction of cDNA containing complete coding sequence of the CFTR, only restriction sites relevant to the DNA constructions described below are shown;

Figure 2 depicts plasmid construction of the CFTR cDNA clone pKK-CFTR1;

Figure 3 depicts plasmid construction of the CFTR cDNA clone pKK-CFTR2;

Figure 4 depicts plasmid construction of the CFTR cDNA clone pSC-CFTR2;

Figure 5 shows a plasmid map of the CFTR cDNA clone pSC-CFTR2;

Figure 6 shows the DNA sequence of synthetic DNAs used for insertion of an intron into the CFTR cDNA sequence, with the relevant restriction endonuclease sites and nucleotide positions noted;

Figures 7A and 7B depict plasmid construction of the CFTR cDNA clone pKK-CFTR3;

Figure 8 shows a plasmid map of the CFTR cDNA pKK-CFTR3 containing an intron between nucleotides 1716 and 1717;

Figure 9 shows treatment of CFTR with glycosidases;

Figures 10A and 10B show an analysis of CFTR expressed from COS-7 transfected cells;

Figures 11A and 11B show pulse-chase labeling of wild type and ΔF508 mutant CFTR in COS-7 transfected cells;

Figures 12A-12D show immunolocalization of wild type and ΔF508 mutant CFTR; and COS-7 cells transfected with pMT-CFTR or pMT-CFTR-ΔF508;

Figure 13 shows an analysis of mutant forms of CFTR;

Figure 14 shows a map of the first generation adenovirus based vector encoding CFTR (Ad2/CFTR-1);

Figure 15 shows the plasmid construction of the Ad2/CFTR-1 vector;

Figure 16 shows an example of UV fluorescence from an agarose gel electrophoresis of products of nested RT-PCR from lung homogenates of cotton rats which received Ad2/CFTR-1. The gel demonstrates that the homogenates were positive for virally-encoded CFTR mRNA;

Figure 17 shows an example of UV fluorescence from an agarose gel electrophoresis of products of nested RT-PCR from organ homogenates of cotton rats. The gel demonstrates that all organs of the infected rats were negative for Ad2/CFTR with the exception of the small bowel;

Figures 18A and 18B show differential cell analyses of bronchoalveolar lavage specimens from control and infected rats. These data demonstrate that none of the rats treated with Ad2/CFTR-1 had a change in the total or differential white blood cell count 4, 10, and 14 days after infection (Figure 18A) and 3, 7, and 14 days after infection (Figure 18B);

Figure 19 shows hematoxilyn and eosin stained sections of cotton rat tracheas from both treated and control rats sacrificed at different time points after infection with Ad2/CFTR-1. The sections demonstrate that there were no observable differences between the treated and control rats;

Figures 20A and 20B show examples of UV fluorescence from an agarose gel electrophoresis, stained with ethidium bromide, of products of RT-PCR from nasal brushings of Rhesus monkeys after application of Ad2/CFTR-1 or Ad2/β-Gal;

Figure 21 shows lights microscopy and immunocytochemistry from monkey nasal brushings. The microscopy revealed that there was a positive reaction when nasal epithelial cells from monkeys exposed to Ad2/CFTR-1 were stained with antibodies to CFTR;

Figure 22 shows immunocytochemistry of monkey nasal turbinate biopsies. This microscopy reveals increased immunofluorescence at the apical membrane of the surface epithelium from biopsies obtained from monkeys treated with Ad2/CFTR-1 over that seen at the apical membrane of the surface epithelium from biopsies obtained from control monkeys;

Figures 23A-23D show serum antibody titers in Rhesus monkeys after three vector administrations. These graphs demonstrate that all three monkeys treated with Ad2/CFTR-1 developed antibodies against adenovirus;

Figure 24 shows hematoxilyn and eosin stained sections from monkey medial turbinate biopsies. These sections demonstrate that turbinate biopsy specimens from control monkeys could not be differentiated from those from monkeys treated with Ad2/CFTR-1 when reviewed by an independent pathologist;

Figures 25A-25I show photomicrographs of human nasal mucosa immediately before, during, and after Ad2/CFTR-1 application. These photomicrographs demonstrate that inspection of the nasal mucosa showed mild to moderate erythema, edema, and exudate in patients treated with Ad2/CFTR-1 (Figures 25A-25C) and in control patients (Figures 25G-25I). These changes were probably due to local anesthesia and vasocontriction because when an additional patient was exposed to Ad2/CFTR in a method which did not require the use of local anesthesia or vasoconstriction, there were no symptoms and the nasal mucosa appeared normal (Figures 25D-25F);

Figure 26 shows a photomicrograph of a hematoxilyn and eosin stained biopsy of human nasal mucosa obtained from the third patient three days after Ad2/CFTR-1 administration. This section shows a morphology consistent with CF, i.e., a thickened basement membrane and occasional morphonuclear cells in the submucosa, but no abnormalities that could be attributed to the adenovirus vector;

Figure 27 shows transepithelial voltage (V_t) across the nasal epithelium of a normal human subject. Amiloride (μM) and terbutaline (μM) were perfused onto the mucosal surface beginning at the times indicated. Under basal conditions (V_t) was electrically negative. Perfusion of amiloride onto the mucosal surface inhibited (V_t) by blocking apical Na^+ channels;

Figures 28A and 28B show transepithelial voltage (V_t) across the nasal epithelium of normal human subjects (Figure 28A) and patients with CF (Figure 28B). Values were obtained under basal conditions, during perfusion with amiloride (μM), and during perfusion of amiloride plus terbutaline (μM) onto the mucosal surface. Data are from seven normal subjects and nine patients with CF. In patients with CF, (V_t) was more electrically negative than in normal subjects (Figure 28B). Amiloride inhibited (V_t) in CF patients, as it did in normal subjects. However, V_t failed to hyperpolarize when terbutaline was perfused onto the epithelium in the presence of amiloride. Instead, (V_t) either did not change or became less negative, a result very different from that observed in normal subjects;

Figures 29A and 29B show transepithelial voltage (V_t) across the nasal epithelium of a third patient before (Figure 29A) and after (Figure 29B) administration of approximately 25 MOI of Ad2/CFTR-1. Amiloride and terbutaline were perfused onto the mucosal surface beginning at the times indicated. Figure 29A shows an example from the third patient before treatment. Figure 29B shows that in contrast to the response before Ad2/CFTR-1 was applied, after virus replication, in the presence of amiloride, terbutaline stimulated V_t;

Figures 30A-30F show the time of course changes in transepithelial electrical properties before and after administration of Ad2/CFTR-1. Figures 30A and 30B are from the first patient who received approximately 1 MOI; Figures 30C and 30D are from the second patient who received approximately 3 MOI; and Figures 30E and 30F are from the third patient who received approximately 25 MOI. Figures 30A, 30C, and 30E show values of basal transeptithelial voltage (V_t) and Figures 30B, 30D, and 30F show the change in transepithelial voltage (ΔV_t) following perfusion of terbutaline in the presence of amiloride. Day zero indicates the day of Ad2/CFTR-1 administration. Figures 30A, 30C, and 30E show the time course of changes in basal V_t for all three patients. The decrease in basal V_t suggests that application of Ad2/CFTR-1 corrected the CF electrolyte transport defect in nasal epithelium of all three patients. Additional evidence came from an examination of the response to terbutaline. Figures 30B, 30D, and 30F show the time course of the response. These data indicate that Ad2/CFTR-1 corrected the CF defect in Cl^- transport;

Figure 31 shows the time course of changes in transepithelial electrical properties before and after administration of saline instead of Ad2/CFTR-1 to CF patients. Day zero indicates the time of mock administration. The top graph shows basal transepithelial voltage (V_t) and the bottom graph shows the change in transepithelial voltage following perfusion with terbutaline in the presence of amiloride (ΔV_t). Closed symbols are data from two patients that received local anesthetic/vasoconstriction and placement of the applicator for thirty minutes. Open symbol is data from a patient that received local anesthetic/vasoconstriction, but not placement of the applicator. Symptomatic changes and physical findings were the same as those observed in CF patients treated with a similar administration procedure and Ad2/CFTR-1;

Figure 32 shows a map of the second generation adenovirus based vector, PAV;

Figure 33 shows the plasmid construction of a second generation adenoviral vector 6 (Ad E4 ORF6);

Figure 34 is a schematic of Ad2-ORF6/PGK-CFTR which differs from Ad2/CFTR in that the latter utilized the endogenous Ela promoter, had no poly A addition signal directly downstream of CFTR and retained an intact E4 region;

Figure 35 shows short-circuit currents from human CF nasal polyp epithelial cells infected with Ad2-ORF6/PGK-CFTR at multiplicities of 0.3, 3, and 50. At the indicated times: (1) 10 μM amiloride, (2) cAMP agonists (10 μM forskolin and 100 μM IBMX, and (3) 1 mM diphenylamine-2-carboxylate were added to the mucosal solution;

Figures 36A-36D show immunocytochemistry of nasal brushings by laser scanning microscopy of the Rhesus monkey C, before infection (36A) and on 7 days (36B); 24 (36C); and 38 (36D) after the first infection with Ad2-ORF6/PGK-CFTR;

Figures 37A-37D show immunocytochemistry of nasal brushings by laser scanning microscopy of Rhesus monkey D, before infection (37A) and on days 7 (37B); 24 (37C); and 48 (37D) after the first infection with Ad2-ORF6/PGK-CFTR;

Figures 38A-38D show immunocytochemistry of nasal brushings by laser scanning microscopy of the Rhesus monkey E, before infection (38A) and on days 7 (38B); 24 (38C); and 48 (38D) after the first infection with Ad2-ORF6/PGK-CFTR;

Figures 39A-39C show summaries of the clinical signs (or lack thereof) of infection with Ad2-ORF6/PGK-CFTR;

Figures 40A-40C shows a summary of blood counts, sedimentation rate, and clinical chemistries after infection with Ad2-ORF6/PGK-CFTR for monkeys C, D, and E. There was no evidence of a systemic inflammatory response or other abnormalities of the clinical chemistries;

Figure 41 shows summaries of white blood cells counts in monkeys C, D, and E after infection with Ad2-ORF6/PGK-CFTR. These date indictate that the administration of Ad2-ORF6/PGK-CFTR caused no change in the distribution and number of inflammatory cells at any of the time points following viral administration;

Figure 42 shows histology of submucosal biopsy performed on Rhesus monkey C.on day 4 after the second viral instillation of Ad2-ORF6/PGK-CFTR. Hematoxylin and eosin stain revealed no evidence of inflammation or cytopathic changes;

Figure 43 shows histology of submucosal biopsy performed on Rhesus monkey D on day 11 after the second viral instillation of Ad2-ORF6/PGK-CFTR. Hematoxylin and eosin stain revealed no evidence of inflammation or cytopathic changes;

Figure 44 shows histology of submucosal biopsy performed on Rhesus monkey E on day 18 after the second viral instillation of Ad2-ORF6/PGK-CFTR. Hematoxylin and eosin stain revealed no evidence of inflammation or cytopathic changes; and

Figures 45A-45C show antibody titers to adenovirus prior to and after the first and second administrations of Ad2-ORF6/PGK-CFTR. Prior to administration of Ad2-ORF6/PGK-

Nucleotide Sequence Analysis (cont.)

```
 8701 TCCGCGTAGG CGCTCGTTGG TCCAGCAGAG GCGGCCGCCC TTGCGCGAAC AGAATGGCGG
 8761 TAGTGGGTCT AGCTGCGTCT CGTCCGGGGG GTCTGCGTCC ACGGTAAAGA CCCCGGGCAG
 8821 CAGGCGCGCG TCGAAGTAGT CTATCTTGCA TCCTTGCAAG TCTAGCGCCT GCTGCCATGC
 8881 GCGGGCGGCA AGCGCGCGCT CGTATGGGTT GAGTGGGGGA CCCCATGGCA TGGGGTGGGT
 8941 GAGCGCGGAG GCGTACATGC CGCAAATGTC GTAAACGTAG AGGGGCTCTC TGAGTATTCC
 9001 AAGATATGTA GGGTAGCATC TTCCACCGCG GATGCTGGCG CGCACGTAAT CGTATAGTTC
 9061 GTGCGAGGGA GCGAGGAGGT CGGGACCGAG GTTGCTACGG GCGGGCTGCT CTGCTCGGAA
 9121 GACTATCTGC CTGAAGATGG CATGTGAGTT GGATGATATG GTTGGACGCT GGAAGACGTT
 9181 GAAGCTGGCG TCTGTGAGAC CTACCGCGTC ACGCACGAAG GAGGCGTAGG AGTCGCGCAG
 9241 CTTGTTGACC AGCTCGGCGG TGACCTGCAC GTCTAGGGCG CAGTAGTCCA GGGTTTCCTT
 9301 GATGATGTCA TACTTATCCT GTCCCTTTTT TTTCCACAGC TCGCGGTTGA GGACAAACTC
 9361 TTCGCGGTCT TTCCAGTACT CTTGGATCGG AAACCCGTCG GCCTCCGAAC GGTAAGAGCC
 9421 TAGCATGTAG AACTGGTTGA CGGCCTGGTA GGCGCAGCAT CCCTTTTCTA CGGGTAGCGC
 9481 GTATGCCTGC GCGGCCTTCC GGAGCGAGGT GTGGGTGAGC GCAAAGGTGT CCCTGACCAT
 9541 GACTTTGAGG TACTGGTATT TGAAGTCAGT GTCGTCGCAT CCGCCCTGCT CCCAGAGCAA
 9601 AAAGTCCGTG CGCTTTTTGG AACGCGGGTT TGGCAGGGCG AAGGTGACAT CGTTGAAAAG
 9661 TATCTTTCCC GCGCGAGGCA TAAAGTTGCG TGTGATGCGG AAGGGTCCCG GCACCTCGGA
 9721 ACGGTTGTTA ATTACCTGGG CGGCGAGCAC GATCTCGTCG AAGCCGTTGA TGTTGTGGCC
 9781 CACGATGTAA AGTTCCAAGA AGCGCGGGGT GCCCTTGATG GAGGGCAATT TTTTAAGTTC
 9841 CTCGTAGGTG AGCTCCTCAG GGGAGCTGAG CCCGTGTTCT GACAGGGCCC AGTCTGCAAG
 9901 ATGAGGGTTG GAAGCGACGA ATGAGCTCCA CAGGTCACGG GCCATTAGCA TTTGCAGGTG
 9961 GTCGCGAAAG GTCCTAAACT GGCGACCTAT GGCCATTTTT TCTGGGGTGA TGCAGTAGAA
10021 GGTAAGCGGG TCTTGTTCCC AGCGGTCCCA TCCAAGGTCC ACGGCTAGGT CTCGCGCGGC
10081 GGTCACCAGA GGCTCATCTC CGCCGAACTT CATAACCAGC ATGAAGGGCA CGAGCTGCTT
10141 CCCAAAGGCC CCCATCCAAG TATAGGTCTC TACATCGTAG GTGACAAAGA GACGCTCGGT
10201 GCGAGGATGC GAGCCGATCG GGAAGAACTG GATCTCCCGC CACCAGTTGG AGGAGTGGCT
10261 GTTGATGTGG TGAAAGTAGA AGTCCCTGCG ACGGGCCGAA CACTCGTGCT GGCTTTTGTA
10321 AAAACGTGCG CAGTACTGGC AGCGGTGCAC GGGCTGTACA TCCTGCACGA GGTTGACCTG
10381 ACGACCGCGC ACAAGGAAGC AGAGTGGGAA TTTGAGCCCC TCGCCTGGCG GGTTTGGCTG
10441 GTGGTCTTCT ACTTCGGCTG CTTGTCCTTG ACCGTCTGGC TGCTCGAGGG GAGTTATGGT
10501 GGATCGGACC ACCACGCCGC GCGAGCCCAA AGTCCAGATG TCCGCGCGCG GCGGTCGGAG
10561 CTTGATGACA ACATCGCGCA GATGGGAGCT GTCCATGGTC TGGAGCTCCC GCGGCGACAG
10621 GTCAGGCGGG AGCTCCTGCA GGTTTACCTC GCATAGCCGG GTCAGGGCGC GGGCTAGGTC
10681 CAGGTGATAC CTGATTTCCA GGGGCTGGTT GGTGGCGGCG TCGATGACTT GCAAGAGGCC
10741 GCATCCCCGC GGCGCGACTA CGGTACCGCG CGGCGGGCGG TGGGCCGCGG GGGTGTCCTT
10801 GGATGATGCA TCTAAAAGCG GTGACGCGGC CGGGCCCCCG GAGGTAGGGG GGGCTCGGGA
10861 CCCGCCGGGA GAGGGGGCAG GGGCACGTCG GCGCCGCGCG CGGGCAGGAG CTGGTGCTGC
10921 GCGCGGAGGT TGCTGGCGAA CGCGACGACG CGGCGGTTGA TCTCCTGAAT CTGGCGCCTC
10981 TGCGTGAAGA CGACGGGCCC GGTGAGCTTG AACCTGAAAG AGAGTTCGAC AGAATCAATT
11041 TCGGTGTCGT TGACGGCGGC CTGGCGCAAA ATCTCCTGCA CGTCTCCTGA GTTGTCTTGA
11101 TAGGCGATTT CGGCCATGAA CTGCTCGATC TCTTCCTCCT GGAGATCTCC GCGTCCGGCT
11161 CGCTCCACGG TGGCGGCGAG GTCGTTGGAG ATGCGGGCCA TGAGCTGCGA GAAGGCGTTG
11221 AGGCCTCCCT CGTTCCAGAC GCGGCTGTAG ACCACGCCCC CTTCGGCATC GCGGGCGCGC
11281 ATGACCACCT GCGCGAGATT GAGCTCCACG TGCCGGGCGA AGACGGCGTA GTTTCGCAGG
11341 CGCTGAAAGA GGTAGTTGAG GGTGGTGGCG GTGTGTTCTG CCACGAAGAA GTACATAACC
11401 CAGCGTCGCA ACGTGGATTC GTTGATATCC CCCAAGGCCT CAAGGCGCTC CATGGCCTCG
11461 TAGAAGTCCA CGGCGAAGTT GAAAAACTGG GAGTTGCGCG CGGACACGGT TAACTCCTCC
11521 TCCAGAAGAC GGATGAGCTC GGCGACAGTG TCGCGCACCT CGCGCTCAAA GGCTACAGGG
11581 GCCTCTTCTT CTTCAATCTC CTCTTCCATA AGGGCCTCCC CTTCTTCTTC TTCTTCTGGC
11641 GGCGGTGGGG GAGGGGGGAC ACGGCGGCGA CGACGGCGCA CCGGGAGGCG GTCGACAAAG
11701 CGCTCGATCA TCTCCCCGCG GCGACGGCGC ATGGTCTCGG TGACGGCGCG GCCGTTCTCG
11761 CGGGGGCGCA GTTGGAAGAC GCCGCCCGTC ATGTCCCGGT TATGGGTTGG CGGGGGGCTG
11821 CCGTGCGGCA GGGATACGGC GCTAACGATG CATCTCAACA ATTGTTGTGT AGGTACTCCG
11881 CCACCGAGGG ACCTGAGCGA GTCCGCATCG ACCGGATCGG AAAACCTCTC GAGAAAGGCG
11941 TCTAACCAGT CACAGTCGCA AGGTAGGCTG AGCACCGTGG CGGGCGGCAG CGGGTGGCGG
12001 TCGGGGTTGT TTCTGGCGGA GGTGCTGCTG ATGATGTAAT TAAAGTAGGC GGTCTTGAGA
12061 CGGCGGATGG TCGACAGAAG CACCATGTCC TTGGGTCCGC CCTGCTGAAT GCGCAGGCGC
```

Nucleotide Sequence Analysis (cont.)

12121 TCGGCCATGC CCCAGGCTTC GTTTTGACAT CGGCGCAGGT CTTTGTAGTA GTCTTGCATG
12181 AGCCTTTCTA CCGGCACTTC TTCTTCTCCT TCCTCTTGTC CTGCATCTCT TGCATCTATC
12241 GCTACGGCGG CGGCGGAGTT TGGCCGTAGG TGGCGCCCTC TTCCTCCCAT GCGTGTGACC
12301 CCGAAGCCCC TCATCGGCTG AAGCAGGGCC AGGTCGGCGA CAACGCGCTC GGCTAATATG
12361 GCCTGCTGCA CCTGCGTGAG GGTAGACTGG AAGTCATCCA TGTCCACAAA GCGGTGGTAT
12421 GCGCCCGTGT TGATGGTGTA AGTGCAGTTG GCCATAACGG ACCAGTTAAC GGTCTGGTGA
12481 CCCGGCTGCG AGAGCTCGGT GTACCTGAGA CGCGAGTAAG CCCTTGAGTC AAAGACGTAG
12541 TCGTTGCAAG TCCGCACCAG GTACTGATAT CCCACCAAAA AGTGCGGCGG CGGCTGGCGG
12601 TAGAGGGGCC AGCGTAGGGT GGCCGGGGCT CCGGGGCGGA GGTCTTCCAA CATAAGGCGA
12661 TGATATCCGT AGATGTACCT GGACATCCAG GTGATGCCGG CGGCGGTGGT GGAGGCGCGC
12721 GGAAAGTCGC GGACGCGGTT CCAGATGTTG CGCAGCGGCA AAAAGTGCTC CATGGTCGGG
12781 ACGCTCTGGC CGGTGAGGCG TGCGCAGTCG TTGACGCTCT AGACCGTGCA AAAGGAGAGC
12841 CTGTAAGCGG GCACTCTTCC GTGGTCTGGT GGATAAATTC GCAAGGGTAT CATGGCGGAC
12901 GACCGGGGTT CGAACCCCGG ATCGGGCCGT CCGCCGTGAT CCATGCGGTT ACCGCCCGCG
12961 TGTCGAACCC AGGTGTGCGA CGTCAGACAA CGGGGGAGCG CTCCTTTTGG CTTCCTTCCA
13021 GGCGCGGCGG CTGCTGCGCT AGCTTTTTTG GCCACTGGCC GCGCGCGCGC TAAGCGGTTA
13081 GGCTGGAAAG CGAAAGCATT AAGTGGCTCG CTCCCTGTAG CCGGAGGGTT ATTTTCCAAG
13141 GGTTGAGTCG CAGGACCCCC GGTTCGAGTC TCGGGCCGGC CGGACTGCGG CGAACGGGGG
13201 TTTGCCTCCC CGTCATGCAA GACCCCGCTT GCAAATTCCT CCGGAAACAG GGACGAGCCC
13261 CTTTTTTGCT TTTCCCAGAT GCATCCGGTG CTGCGGCAGA TGCGCCCCCC TCCTCAGCAG
13321 CGGCAAGAGC AAGAGCAGCG GCAGACATGC AGGGCACCCT CCCCTTCTCC TACCGCGTCA
13381 GGAGGGGCAA CATCCGCGGC TGACGCGGCG GCAGATGGTG ATTACGAACC CCCGCGGCGC
13441 CGGGCCGGGC ACTACCTGGA CTTGGAGGAG GCGCAGGGCC TGGCGCGGCT AGGAGCGCCC
13501 TCTCCTGAGC GACACCCAAG GGTGCAGCTG AAGCGTGACA CGCGCGAGGC GTACGTGCCG
13561 CGGCAGAACC TGTTTCGCGA CCGCGAGGGA GAGGAGCCCG AGGAGATGCG GGATCGAAAG
13621 TTCCACGCAG GCGCGAGTT GCGGCATGGC CTGAACCGCG AGCGGTTGCT GCGCGAGGAG
13681 GACTTTGAGC CGAACGCGCG GACCGGGATT AGTCCCGCGC GCGCACACGT GGCGGCCGCC
13741 GACCTGGTAA CCGCGTACGA GCAGACGGTG AACCAGGAGA TTAACTTTCA AAAAAGCTTT
13801 AACAACCACG TGCGCACGCT TGTGGCGCGC GAGGAGGTGG CTATAGGACT GATGCATCTG
13861 TGGGACTTTG TAAGCCGCCT GGAGCAAAAC CCAAATAGCA AGCCGCTCAT GGCGCAGCTG
13921 TTCCTTATAG TGCAGCACAG CAGGGACAAC GAGGCATTCA GGGATGCGCT GCTAAACATA
13981 GTAGAGCCCG AGGGCCGCTG GCTGCTCGAT TTGATAAACA TTCTGCAGAG CATAGTGGTG
14041 CAGGAGCGCA GCTTGAGCCT GGCTGACAAG GTGGCCGCCA TTAACTATTC CATGCTCAGT
14101 CTGGGCAAGT TTTACGCCCG CAAGATATAC CATACCCCTT ACGTTCCCAT AGACAAGGAG
14161 GTAAAGATCG AGGGGTTCTA CATGCGCATG GCGTTGAAGG TGCTTACCTT GAGCGACGAC
14221 CTGGGCGTTT ATCGCAACGA GCGCATCCAC AAGGCCGTGA GCGTGAGCCG GCGGCGCGAG
14281 CTCAGCGACC GCGAGCTGAT GCACAGCCTG CAAAGGGCCC TGGCTGGCAC GGGCAGCGGC
14341 GATAGAGAGG CCGAGTCCTA CTTTGACGCG GGCGCTGACC TGCGCTGGGC CCCAAGCCGA
14401 CGCGCCCTGG AGGCAGCTGG GGCCGGACCT GGGCTGGCGG TGGCACCCGC GCGCGCTGGC
14461 AACGTCGGCG GCGTGGAGGA ATATGACGAG GACGATGAGT ACGAGCCAGA GGACGGCGAG
14521 TACTAAGCGG TGATGTTTCT GATCAGATGA TGCAAGACGC AACGGACCCG GCGGTGCGGG
14581 CGGCGCTGCA GAGCCAGCCG TCGGGCCTTA ACTCCACGGA CGACTGGCGC CAGGTCATGG
14641 ACCGCATCAT GTCGCTGACT GCGCGTAACC CTGACGCGTT CCGGCAGCAG CCGCAGGCCA
14701 ACCGGCTCTC CGCAATTCTG GAAGCGGTGG TCCCGGCGCG CGCAAACCCC ACGCACGAGA
14761 AGGTGCTGGC GATCGTAAAC GCGCTGGCCG AAAACAGGGC CATCCGGCCC GATGAGGCCG
14821 GCCTGGTCTA CGACGCGCTG CTTCAGCGCG TGGCTCGTTA CAACAGCGGC AACGTGCAGA
14881 CCAACCTGGA CCGGCTGGTG GGGGATGTGC GCGAGGCCGT GGCGCAGCGT GAGCGCGCGC
14941 AGCAGCAGGG CAACCTGGGC TCCATGGTTG CACTAAACGC CTTCCTGAGT ACACAGCCCG
15001 CCAACGTGCC GCGGGGACAG GAGGACTACA CCAACTTTGT GAGCGCACTG CGGCTAATGG
15061 TGACTGAGAC ACCGCAAAGT GAGGTGTACC AGTCCGGGCC AGACTATTTT TTCCAGACCA
15121 GTAGACAAGG CCTGCAGACC GTAAACCTGA GCCAGGCTTT CAAGAACTTG CAGGGGCTGT
15181 GGGGGGTGCG GGCTCCCACA GCCGACCGCG CGACCGTGTC TAGCTTGCTG ACGCCCAACT
15241 CGCGCCTGTT GCTGCTGCTA ATAGCGCCCT TCACGGACAG TGGCAGCGTG TCCCGGGACA
15301 CATACCTAGG TCACTTGCTG ACACTGTACC GCGAGGCCAT AGGTCAGGCG CATGTGGACG
15361 AGCATACTTT CCAGGAGATT ACAAGTGTCA GCCGCCGCCT GGGGCAGGAG GACACCGCCA
15421 GCCTGGAGGC AACCCTGAAC TACCTGCTGA CCAACCGGCG GCAGAAGATC CCCTCGTTGC
15481 ACAGTTTAAA CAGCGAGGAG GAGCGCATCT TGCGCTATGT GCAGCAGAGC GTGAGCCTTA

Nucleotide Sequence Analysis (cont.)

15541 ACCTGATGCG CGACGGGGTA ACGCCCAGCG TGGCGCTGGA CATGACCGCG CGCAACATGG
15601 AACCGGGCAT GTATGCCTCA AACCGGCCGT TTATCAATCG CCTAATGGAC TACTTGCATC
15661 GCGCGGCCGC CGTGAACCCC GAGTATTTCA CCAATGCCAT CTTGAACCCG CACTGGCTAC
15721 CGCCCCCTGG TTTCTACACC GGGGGATTTG AGGTGCCCGA GGGTAACGAT GGATTCCTCT
15781 GGGACGACAT AGACGACAGC GTGTTTTCCC CGCAACCGCA GACCCTGCTA GAGTTGCAAC
15841 AGCGCGAGCA GGCAGAGGCG GCGCTGCGAA AGGAAAGCTT CCGCAGGCCA AGCAGCTTGT
15901 CCGATCTAGG CGCTGCGGCC CCGCGGTCAG ATGCGAGTAG CCCATTTCCA AGCTTGATAG
15961 GGTCTTTTAC CAGCACTCGC ACCACCCGCC CGCGCCTGCT GGGCGAGGAG GAGTACCTAA
16021 ACAACTCGCT GCTGCAGCCG CAGCGCGAAA AGAACCTGCC TCCGGCATTT CCCAACAACG
16081 GGATAGAGAG CCTAGTGGAC AAGATGAGTA GATGGAAGAC GTATGCGCAG GAGCACAGGG
16141 ATGTGCCCGG CCCGCGCCCG CCCACCCGTC GTCAAAGGCA CGACCGTCAG CGGGGTCTGG
16201 TGTGGGAGGA CGATGACTCG GCAGACGACA GCAGCGTCCT GGATTTGGGA GGGAGTGGCA
16261 ACCCGTTTGC GCACCTTCGC CCCAGGCTGG GGAGAATGTT TTAAAAAAAA AAAAAAAAAG
16321 CATGATGCAA AATAAAAAAC TCACCAAGGC CATGGCACCG AGCGTTGGTT TTCTTGTATT
16381 CCCCTTAGTA TGCAGCGCGC GGCGATGTAT GAGGAAGGTC CTCCTCCCTC CTACGAGAGC
16441 GTGGTGAGCG CGGCGCCAGT GGCGGCGGCG CTGGGTTCCC CCTTCGATGC TCCCCTGGAC
16501 CCGCCGTTTG TGCCTCCGCG GTACCTGCGG CCTACCGGGG GGAGAAACAG CATCCGTTAC
16561 TCTGAGTTGG CACCCCTATT CGACACCACC CGTGTGTACC TTGTGGACAA CAAGTCAACG
16621 GATGTGGCAT CCCTGAACTA CCAGAACGAC CACAGCAACT TTCTAACCAC GGTCATTCAA
16681 AACAATGACT ACAGCCCGGG GGAGGCAAGC ACACAGACCA TCAATCTTGA CGACCGTTCG
16741 CACTGGGGCG GCGACCTGAA AACCATCCTG CATACCAACA TGCCAAATGT GAACGAGTTC
16801 ATGTTTACCA ATAAGTTTAA GGCCGGGGTG ATGGTGTCGC GCTCGCTTAC TAAGGACAAA
16861 CAGGTGGAGC TGAAATATGA GTGGGTGGAG TTCACGCTGC CCGAGGGCAA CTACTCCGAG
16921 ACCATGACCA TAGACCTTAT GAACAACGCG ATCGTGGAGC ACTACTTGAA AGTGGGCAGG
16981 CAGAACGGGG TTCTGGAAAG CGACATCGGG GTAAAGTTTG ACACCCGCAA CTTCAGACTG
17041 GGGTTTGACC CAGTCACTGG TCTTGTCATG CCTGGGGTAT ATACAAACGA AGCCTTCCAT
17101 CCAGACATCA TTTTGCTGCC AGGATGCGGG GTGGACTTCA CCCACAGCCG CCTGAGCAAC
17161 TTGTTGGGCA TCCGCAAGCG GCAACCCTTC CAGGAGGGCT TTAGGATCAC CTACGATGAC
17221 CTGGAGGGTG GTAACATTCC CGCACTGTTG GATGTGGACG CCTACCAGGC AAGCTTAAAA
17281 GATGACACCG AACAGGGCGG GGATGGCGCA GGCGGCGGCA ACAACAGTGG CAGCGGCGCG
17341 GAAGAGAACT CCAACGCGGC AGCCGCGGCA ATGCAGCCGC TGGAGGACAT GAACGATCAT
17401 GCCATTCGCG GCGACACCTT TGCCACACGG GCGGAGGAGA AGCGCGCTGA GGCCGAGGCA
17461 GCGGCAGAAG CTGCCGCCCC CGCTGCGCAA CCCGAGGTCG AGAAGCCTCA GAAGAAACCG
17521 GTGATCAAAC CCCTGACAGA GGACAGCAAG AAACGCAGTT ACAACCTAAT AAGCAATGAC
17581 AGCACCTTCA CCCAGTACCG CAGCTGGTAC CTTGCATACA ACTACGGCGA CCCTCAGACC
17641 GGGATCCGCT CATGGACCCT CCTTTGCACT CCTGACGTAA CCTGCGGCTC GGAGCAGGTC
17701 TACTGGTCGT TGCCAGACAT GATGCAAGAC CCCGTGACCT TCCGCTCCAC GAGCCAGATC
17761 AGCAACTTTC CGGTGGTGGG CGCCGAGCTG TTGCCCGTGC ACTCCAAGAG CTTCTACAAC
17821 GACCAGGCCG TCTACTCCCA GCTCATCCGC CAGTTTACCT CTCTGACCCA CGTGTTCAAT
17881 CGCTTTCCCG AGAACCAGAT TTTGCGCGCC CCGCCAGCCC CCACCATCAC CACCGTCAGT
17941 GAAAACGTTC CTGCTCTCAC AGATCACGGG ACGCTACCGC TGCGCAACAG CATCGGAGGA
18001 GTCCAGCGAG TGACCATTAC TGACGCCAGA CGCCGCACCT GCCCCTACGT TTACAAGGCC
18061 CTGGGCATAG TCTCGCCGCG CGTCCTATCG AGCCGCACTT TTTGAGCAAA CATGTCCATC
18121 CTTATATCGC CCAGCAATAA CACAGGCTGG GGCCTGCGCT TCCCAAGCAA GATGTTTGGC
18181 GGGGCAAAGA AGCGCTCCGA CCAACACCCA GTGCGCGTGC GCGGGCACTA CCGCGCGCCC
18241 TGGGGCGCGC ACAAACGAGC CCGCACTGGG CGCACCACCG TCGATGACGC CATTGACGCG
18301 GTGGTGGAGG AGGCGCGCAA CTACACGCCC ACGCCGCCAC CAGTGTCCAC AGTGGACGCG
18361 GCCATTCAGA CCGTGGTGCG CGGAGCCCGG CGTTATGCTA AAATGAAGAG ACGGCGGAGG
18421 CGCGTAGCAC GTCGCCACCG CCGCCGACCC GGCACTGCCG CCCAACGCGC GGCGGCGGCC
18481 CTGCTTAACC GCGCACGTCG CACCGGCCGA CGGGCGGCCA TGCGGGCCGC TCGAAGGCTG
18541 GCCGCGGGTA TTGTCACTGT GCCCCCCAGG TCCAGGCGAC GAGCGGCCGC CGCAGCAGCC
18601 GCGGCCATTA GTGCTATGAC TCAGGGTCGC AGGGGCAACG TGTACTGGGT GCGCGACTCG
18661 GTTAGCGGCC TGCGCGTGCC CGTGCGCACC CGCCCCCCGC GCAACTAGAT TGCAAGAAAA
18721 AACTACTTAG ACTCGTACTG TTGTATGTAT CCAGCGGCGG CGGCGCGCAA CGAAGCTATG
18781 TCCAAGCGCA AAATCAAAGA AGAGATGCTC CAGGTCATCG CGCCGGAGAT CTATGGCCCC
18841 CCGAAGAAGG AAGAGCAGGA TTACAAGCCC CGAAAGCTAA AGCGGGTCAA AAAGAAAAAG
18901 AAAGATGATG ATGATGATGA ACTTGACGAC GAGGTGGAAC TGCTGCACGC AACCGCGCCC.

Nucleotide Sequence Analysis (cont.)

```
18961 AGGCGGCGGG TACAGTGGAA AGGTCGACGC GTAAGACGTG TTTTGCGACC CGGCACCACC
19021 GTAGTTTTTA CGCCCGGTGA GCGCTCCACC CGCACCTACA AGCGCGTGTA TGATGAGGTG
19081 TACGGCGACG AGGACCTGCT TGAGCAGGCC AACGAGCGCC TCGGGGAGTT TGCCTACGGA
19141 AAGCGGCATA AGGACATGTT GGCGTTGCCG CTGGACGAGG GCAACCCAAC ACTTAGCCTA
19201 AAGCCCGTGA CACTGCAGCA GGTGCTGCCC ACGCTTGCAC CGTCCGAAGA AAAGCGCGGC
19261 CTAAAGCGCG AGTCTGGTGA CTTGGCACCC ACCGTGCAGC TGATGGTACC CAAGCGCCAG
19321 CGACTGGAAG ATGTCTTGGA AAAAATGACC GTGGAGCCTG GGCTGGAGCC CGAGGTCCGC
19381 GTGCGGCCAA TCAAGCAGGT GGCACCGGGA CTGGGCGTGC AGACCGTGGA CGTTCAGATA
19441 CCCACCACCA GTAGCACTAG TATTGCCACT GCCACAGAGG GCATGGAGAC ACAAACGTCC
19501 CCGGTTGCCT CGGCGGTGGC AGATGCGGCG GTGCAGGCGG CCGCTGCGGC CGCGTCCAAA
19561 ACCTCTAGGG AGGTGCAAAC GGACCCGTGG ATGTTTCGCG TTTCAGCCCC CCGGCGCCCG
19621 CGCCGTTGCA GGAAGTACGG CACCGCCAGC GCACTACTGC CCGAATATGC CCTACATCCT
19681 TCCATCGCGC CTACCCCCGG CTATCGTGCC TACACCTACC GCCCCAGAAG ACGAGCGACT
19741 ACCCGACGCC GAACCACCAC TGGAACCCGC CGCCGCCGTC GCCGTCGCCA GCCCGTGCTG
19801 GCCCCGATTT CCGTGCGCAG GGTGGCTGGC GAAGGAGGCA GGACCCTGGT GCTGCCAACA
19861 GCGCGCTACC ACCCCAGCAT CGTTTAAAAG CCGGTCTTTG TGGTTCTTGC AGATATGGCC
19921 CTCACCTGCC GCCTCCGTTT CCCCGTGCCG GGATTCCGAG GAAGAATGCA CCGTAGGAGG
19981 GGCATGGCCG GCCACGGCCT GACGGGCGGC ATGCGTCGTG CGCACCACCG CGGCGGCGGC
20041 GCGTCGCACC GTCGCATGCG CGGCGGTATC CTGCCCCTCC TTATTCCACT GATCGCCGCG
20101 GCGATTGGCG CCGTGCCCGG AATTGCATCC GTGGCCTTGC AGGCGCAGAG ACACTGATTA
20161 AAAACAAGTT GCATGTGGAA AAATCAAAAT AAAAAGTCTG GAGTCTCACG CTGCTTGGT
20221 CCTGTAACTA TTTTGTAGAA TGGAAGACAT CAACTTTGCG TCTCTGGCCC CGCGACACGG
20281 CTCGCGCCCG TTCATGGGAA ACTGGCAAGA TATCGGCACC AGCAATATGA GCGGTGGCGC
20341 CTTCAGCTGC GGCTCGCTGT GGAGCGGCAT TAAAAATTTC GGTTCCACCA TTAAGAACTA
20401 TGGCAGCAAG GCCTGGAACA GCAGCACAGG CCAGATGCTG AGGGACAAGT TGAAAGAGCA
20461 AAATTTCCAA CAAAAGGTGG TAGATGGCCT GGCCTCTGGC ATTAGCGGGG TGGTGGACCT
20521 GGCCAACCAG GCAGTGCAAA ATAAGATTAA CAGTAAGCTT GATCCCCGCC CTCCCGTAGA
20581 GGAGCCTCCA CCGGCGGTGG AGACAGTGTC TCCAGAGGGG CGTGGCGAAA AGCGTCCGCG
20641 GCCCGACAGG GAAGAAACTC TGGTGAGCGA AATAGATGAG CCTCCCTCGT ACGAGGAGGC
20701 ACTAAAGCAA GGCCTGCCCA CCACCCGTCC CATCGCCGCC ATGGCTACCG GAGTGCTGGG
20761 CCAGCACACA CCTGTAACGC TGGACCTGCC TCCCCCCGCT GACACCCAGC AGAAACCTGT
20821 GCTGCCAGGG CCGTCCGCCG TTGTTGTAAC CCGCCCTAGC CGCCGTCCC TGCGCCGTGC
20881 CGCCAGCGGT CCGCGATGGA TGCGGCCCGT AGCCAGTGGC AACTGGCAAA GCACACTGAA
20941 CAGCATGGTG GGTCTGGGGG TGCAATCCCT GAAGCGCCGA CGATGCTTCT AAATAGCTAA
21001 CGTGTCGTAT GTGTCATGTA TGCGTCCATG TCGCCGCCAG AGGAGCTGCT GAGCCGCCGT
21061 GCGCCCGCTT TCCAAGATGG CTACCCCTTC GATGATGCCG CAGTGGTCTT ACATGCACAT
21121 CTCGGGCCAG GACGCCTCGG AGTACCTGAG CCCCGGGCTG GTGCAGTTTG CCCGCGCCAC
21181 CGAGACGTAC TTCAGCCTGA ATAACAAGTT TAGAAACCCC ACGGTGGCAC CTACGCACGA
21241 CGTAACCACA GACCGGTCCC AGCGTTTGAC GCTGCGGTTC ATCCCTGTGG ACCCGCGAGA
21301 TACCGCGTAC TCGTACAAAG CGCGGTTCAC CCTGGCTGTG GGTGACAACC GTGTGCTTGA
21361 TATGGCTTCC ACGTACTTTG ACATCCGCGG CGTGCTGGAC AGGGGCCTA CTTTTAAGCC
21421 CTACTCCGGC ACTGCCTACA ACGCTCTAGC TCCCAAGGGC GCTCCTAACT CCTGTGAGTG
21481 GGAACAAACC GAAGATAGCG GCCGGGCAGT TGCCGAGGAT GAAGAAGAGG AAGATGAAGA
21541 TGAAGAAGAG GAAGAAGAAG AGCAAAACGC TCGAGATCAG GCTACTAAGA AAACACATGT
21601 CTATGCCCAG GCTCCTTTGT CTGGAGAAAC AATTACAAAA AGCGGGCTAC AAATAGGATC
21661 AGACAATGCA GAAACACAAG CTAAACCTGT ATACGCAGAT CCTTCCTATC AACCAGAACC
21721 TCAAATTGGC GAATCTCAGT GGAACGAAGC TGATGCTAAT GCGGCAGGAG GGAGAGTGCT
21781 TAAAAAAACA ACTCCCATGA AACCATGCTA TGGATCTTAT GCCAGGCCTA CAAATCCTTT
21841 TGGTGGTCAA TCCGTTCTGG TTCCGGATGA AAAAGGGGTG CCTCTTCCAA AGGTTGACTT
21901 GCAATTCTTC TCAAATACTA CCTCTTTGAA CGACCGGCAA GGCAATGCTA CTAAACCAAA
21961 AGTGGTTTTG TACAGTGAAG ATGTAAATAT GGAAACCCCA GACACACATC TGTCTTACAA
22021 ACCTGGAAAA GGTGATGAAA ATTCTAAAGC TATGTTGGGT CAACAATCTA TGCCAAACAG
22081 ACCCAATTAC ATTGCTTTCA GGGACAATTT TATTGGCCTA ATGTATTATA ACAGCACTGG
22141 CAACATGGGT GTTCTTGCTG GTCAGGCATC GCAGCTAAAT GCCGTGGTAG ATTTGCAAGA
22201 CAGAAACACA GAGCTGTCCT ATCAACTCTT GCTTGATTCC ATAGGTGATA GAACCAGATA
22261 TTTTTCTATG TGGAATCAGG CTGTAGACAG CTATGATCCA GATGTTAGAA TCATTGAAAA
22321 CCATGGAACT GAGGATGAAT TGCCAAATTA TTGTTTTCCT CTTGGGGGTA TTGGGGTAAC
```

Nucleotide Sequence Analysis (cont.)

22381 TGACACCTAT CAAGCTATTA AGGCTAATGG CAATGGCTCA GGCGATAATG GAGATACTAC
22441 ATGGACAAAA GATGAAACTT TTGCAACACG TAATGAAATA GGAGTGGGTA ACAACTTTGC
22501 CATGGAAATT AACCTAAATG CCAACCTATG GAGAAATTTC CTTTACTCCA ATATTGCGCT
22561 GTACCTGCCA GACAAGCTAA AATACAACCC CACCAATGTG GAAATATCTG ACAACCCCAA
22621 CACCTACGAC TACATGAACA AGCGAGTGGT GGCTCCCGGG CTTGTAGACT GCTACATTAA
22681 CCTTGGGGCG CGCTGGTCTC TGGACTACAT GGACAACGTT AATCCCTTTA ACCACCACGC
22741 CAATGCGGGC CTCCGTTATC GCTCCATGTT GTTGGGAAAC GGCGGCTACG TGCCCTTTCA
22801 CATTCAGGTG CCCCAAAAGT TTTTTGCCAT TAAAAACCTC CTCCTCCTGC CAGGCTCATA
22861 TACATATGAA TGGAACTTCA GGAAGGATGT TAACATGGTT CTGCAGAGCT CTCTGGGAAA
22921 CGATCTTAGA GTTGACGGGG CTAGCATTAA GTTTGACAGC ATTTGTCTTT ACGCCACCTT
22981 CTTCCCCATG GCCCACAACA CGGCCTCCAC GCTGGAAGCC ATGCTCAGAA ATGACACCAA
23041 CGACCAGTCC TTTAATGACT ACCTTTCCGC CGCCAACATG CTATACCCCA TACCCGCCAA
23101 CGCCACCAAC GTGCCCATCT CCATCCCATC GCGCAACTGG GCAGCATTTC GCGGTTGGGC
23161 CTTCACACGC TTGAAGACAA AGGAAACCCC TTCCCTGGGA TCAGGCTACG ACCCTTACTA
23221 CACCTACTCT GGCTCCATAC CATACCTTGA CGGAACCTTC TATCTTAATC ACACCTTTAA
23281 GAAGGTGGCC ATTACCTTTG ACTCTTCTGT TAGCTGGCCG GGCAACGACC GCCTGCTTAC
23341 TCCCAATGAG TTTGAGATTA AACGCTCAGT TGACGGGGAG GGCTACAACG TAGCTCAGTG
23401 CAACATGACC AAGGACTGGT TCCTGGTGCA GATGTTGGCC AACTACAATA TTGGCTACCA
23461 GGGCTTCTAC ATTCCAGAAA GCTACAAGGA CCGCATGTAC TCGTTCTTCA GAAACTTCCA
23521 GCCCATGAGC CGGCAAGTGG TTGACGATAC TAAATACAAG GAGTATCAGC AGGTTGGAAT
23581 TCTTCACCAG CATAACAACT CAGGATTCGT AGGCTACCTC GCTCCCACCA TGCGCGAGGG
23641 ACAGGCTTAC CCCGCCAACG TGCCCTACCC ACTAATAGGC AAAACCGCGG TTGACAGTAT
23701 TACCGAGAAA AAGTTTCTTT GCGATCGCAC CCTTTGGCGC ATCCCATTCT CCAGTAACTT
23761 TATGTCCATG GCGCACTCA CAGACCTGGG CCAAAACCTT CTCTACGCCA ACTCCGCCCA
23821 CGCGCTAGAC ATGACTTTTG AGGTGGATCC CATGGACGAG CCCACCCTTC TTTATGTTTT
23881 GTTTGAAGTC TTTGACGTGG TCCGTGTGCA CCAGCCGCAC CGCGGCGTCA TCGAGACCGT
23941 GTACCTGCGC ACGCCCTTCT CGGCCGGCAA CGCCACAACA TAAAAGAAGC AAGCAACATC
24001 AACAACAGCT GCCGCCATGG GCTCCAGTGA GCAGGAACTG AAAGCCATTG TCAAAGATCT
24061 TGGTTGTGGG CCATATTTTT TGGGCACCTA TGACAAGCGC TTTCCAGGCT TTGTTTCTCC
24121 ACACAAGCTC GCCTGCGCCA TAGTCAATAC GGCCGGTCGC GAGACTGGGG GCGTACACTG
24181 GATGGCCTTT GCCTGGAACC CGCGCTCAAA AACATGCTAC CTCTTTGAGC CCTTTGGCTT
24241 TTCTGACCAA CGACTCAAGC AGGTTTACCA GTTTGAGTAC GAGTCACTCC TGCGCCGTAG
24301 CGCCATTGCT TCTTCCCCCG ACCGCTGTAT AACGCTGGAA AAGTCCACCC AAAGCGTGCA
24361 GGGGCCCAAC TCGGCCGCCT GTGGACTATT CTGCTGCATG TTTCTCCACG CCTTTGCCAA
24421 CTGGCCCCAA ACTCCCATGG ATCACAACCC CACCATGAAC CTTATTACCG GGGTACCCAA
24481 CTCCATGCTT AACAGTCCCC AGGTACAGCC CACCCTGCGT CGCAACCAGG AACAGCTCTA
24541 CAGCTTCCTG GAGCGCCACT CGCCCTACTT CCGCAGCCAC AGTGCGCAGA TTAGGAGCGC
24601 CACTTCTTTT TGTCACTTGA AAAACATGTA AAAATAATGT ACTAGGAGAC ACTTTCAATA
24661 AAGGCAAATG TTTTTATTTG TACACTCTCG GGTGATTATT TACCCCCCAC CCTTGCCGTC
24721 TGCGCCGTTT AAAAATCAAA GGGGTTCTGC CGCGCATCGC TATGCGCCAC TGGCAGGGAC
24781 ACGTTGCGAT ACTGGTGTTT AGTGCTCCAC TTAAACTCAG GCACAACCAT CGCGGCAGC
24841 TCGGTGAAGT TTTCACTCCA CAGGCTGCGC ACCATCACCA ACGCGTTTAG CAGGTCGGGC
24901 GCCGATATCT TGAAGTCGCA GTTGGGGCCT CCGCCCTGCG CGCGCGAGTT GCGATACACA
24961 GGGTTGCAGC ACTGGAACAC TATCAGCGCC GGGTGGTGCA CGCTGGCCAG CACGCTCTTG
25021 TCGGAGATCA GATCCGCGTC CAGGTCCTCC GCGTTGCTCA GGGCGAACGG AGTCAACTTT
25081 GGTAGCTCCC TTCCCAAAAA GGGTGCATGC CCAGGCTTTG AGTTGCACTC GCACCGTAGT
25141 GGCATCAGAA GGTGACCGTG CCCGGTCTGG GCGTTAGGAT ACAGCGCCTG CATGAAAGCC
25201 TTGATCTGCT TAAAAGCCAC CTGAGCCTTT GCGCCTTCAG AGAAGAACAT GCCGCAAGAC
25261 TTGCCGGAAA ACTGATTGGC CGGACAGGCC GCGTCATGCA CGCAGCACCT TGCGTCGGTG
25321 TTGGAGATCT GCACCACATT TCGGCCCCAC CGGTTCTTCA CGATCTTGGC CTTGCTAGAC
25381 TGCTCCTTCA GCGCGCGCTG CCCGTTTTCG CTCGTCACAT CCATTTCAAT CACGTGCTCC
25441 TTATTTATCA TAATGCTCCC GTGTAGACAC TTAAGCTCGC CTTCGATCTC AGCGCAGCGG
25501 TGCAGCCACA ACGCGCAGCC CGTGGGCTCG TGGTGCTTGT AGGTTACCTC TGCAAACGAC
25561 TGCAGGTACG CCTGCAGGAA TCGCCCCATC ATCGTCACAA AGGTCTTGTT GCTGGTGAAG
25621 GTCAGCTGCA ACCCGCGGTG CTCCTCGTTT AGCCAGGTCT TGCATACGGC CGCCAGAGCT
25681 TCCACTTGGT CAGGCAGTAG CTTGAAGTTT GCCTTTAGAT CGTTATCCAC GTGGTACCTG
25741 TCCATCAACG CGCGCGCAGC CTCCATGCCC TTCTCCCACG CAGACACGAT CGGCAGGCTC

Nucleotide Sequence Analysis (cont.)

25801 AGCGGGTTTA TCACCGTGCT TTCACTTTCC GCTTCACTGG ACTCTTCCTT TTCCTCTTGC
25861 GTCCGCATAC CCCGCGCCAC TGGGTCGTCT TCATTCAGCC GCCGCACCGT GCGCTTACCT
25921 CCCTTGCCGT GCTTGATTAG CACCGGTGGG TTGCTGAAAC CCACCATTTG TAGCGCCACA
25981 TCTTCTCTTT CTTCCTCGCT GTCCAGGATC ACCTCTGGGG ATGGCGGGCG CTCGGCCTTG
26041 GGAGAGGGGC GCTTCTTTTT CTTTTTGCAC GCAATGGCCA AATCCGCCGT CGAGGTCGAT
26101 GGCCGCGGGC TGGGTGTGCG CGGCACCAGC GCATCTTGTG ACGAGTCTTC TTCGTCCTCG
26161 GACTCGAGAC GCCGCCTCAG CCGCTTTTTT GGGGGCGCGC GGGGAGGCGG CGGCGACGGC
26221 GACGGGGACG ACACGTCCTC CATGGTTGGT GGACGTCGCG CCGCACCGCG TCCGCGCTCG
26281 GGGGTGGTTT CGCGCTGCTC CTCTTCCCGA CTGGCCATTT CCTTCTCCTA TAGGCAGAAA
26341 AAGATCATGG AGTCAGTCGA GAAGGAGGAC AGCCTAACCG CCCCCTTTGA GTTCGCCACC
26401 ACCGCCTCCA CCGATGCCGC CAACCGCGCT ACCACCTTCC CCGTCGAGGC ACCCCCGCTT
26461 GAGGAGGAGG AAGTGATTAT CGAGCAGGAC CCAGGTTTTG TAAGCGAAGA CGACGAGGAT
26521 CGCTCAGTAC CAACAGAGGA TAAAAAGCAA GACCAGGACG ACGCAGAGGC AAACGAGGAA
26581 CAAGTCGGGC GGGGGGACCA AAGGCATGGC GACTACCTAG ATGTGGGAGA CGACGTGCTG
26641 TTGAAGCATC TGCAGCGCCA GTCGCGCCATT ATCTGCGACG CGTTGCAAGA GCGCAGCGAT
26701 GTGCCCCTCG CCATAGCGGA TGTCAGCCTT GCCTACGAAC GCCACCTGTT CTCACCGCGC
26761 GTACCCCCCA AACGCCAAGA AAACGGCACA TGCGAGCCCA ACCCCGCGCT CAACTTCTAC
26821 CCCGTATTTG CCGTGCCAGA GGTGCTTGCC ACCTATCACA TCTTTTTCCA AAACTGCAAG
26881 ATACCCCTAT CCTGCCGTGC CAACCGCAGC CGAGCGGACA AGCAGCTGGC CTTGCGGCAG
26941 GGCGCTGTCA TACCTGATAT CGCCTCGCTC GACGAAGTGC CAAAAATCTT TGAGGGTCTT
27001 GGACGCGACG AGAAACGCGC GGCAAACGCT CTGCAACAAG AAAACAGCGA AAATGAAAGT
27061 CACTGTGGAG TGCTGGTGGA ACTTGAGGGT GACAACGCGC GCCTAGCCGT GCTGAAACGC
27121 AGCATCGAGG TCACCCACTT TGCCTACCCG GCACTTAACC TACCCCCCAA GGTTATGAGC
27181 ACAGTCATGA GCGAGCTGAT CGTGCGCCGT GCACGACCCC TGGAGAGGGA TGCAAACTTG
27241 CAAGAACAAA CCGAGGAGGG CCTACCCGCA GTTGGCGATG AGCAGCTGGC GCGCTGGCTT
27301 GAGACGCGCG AGCCTGCCGA CTTGGAGGAC CGACGCAAGC TAATGATGGC CGCAGTGCTT
27361 GTTACCGTGG AGCTTGAGTG CATGCAGCGG TTGTTTGCTG ACCCGGAGAT GCAGCGCAAG
27421 CTAGAGGAAA CGTTGCACTA CACCTTTCGC CAGGGCTACG TGCGCCAGGC CTGCAAAATT
27481 TCCAACGTGG AGCTCTGCAA CCTGGTCTCC TACCTTGGAA TTTTGCACGA AAACCGCCTC
27541 GGGCAAAACG TGCTTCATTC CACGCTCAAG GGCGAGGCGC GCCGCGACTA CGTCCGCGAC
27601 TGCGTTTACT TATTTCTGTG CTACACCTGG CAAACGGCCA TGGGCGTGTG GCAGCAATGC
27661 CTGGAGGAGC GCAACCTAAA GGAGCTGCAG AAGCTGCTAA AGCAAAACTT GAAGGACCTA
27721 TGGACGGCCT TCAACGAGCG CTCCGTGGCC GCGCACCTGG CGGACATTAT CTTCCCCGAA
27781 CGCCTGCTTA AAACCCTGCA ACAGGGTCTG CCAGACTTCA CCAGTCAAAG CATGTTGCAA
27841 AACTTTAGGA ACTTTATCCT AGAGCGTTCA GGAATTCTGC CCGCCACCTG CTGTGCGCTT
27901 CCTAGCGACT TTGTGCCCAT TAAGTACCGT GAATGCCCTC CGCCGCTTTG GGGTCACTGC
27961 TACCTTCTGC AGCTAGCCAA CTACCTTGCC TACCACTCCG ACATCATGGA AGACGTGAGC
28021 GGTGACGGCC TACTGGAGTG TCACTGTCGC TGCAACCTAT GCACCCCGCA CCGCTCCCTG
28081 GTCTGCAATT CGCAACTGCT TAGCGAAAGT CAAATTATCG GTACCTTTGA GCTGCAGGGT
28141 CCCTCGCCTG ACGAAAAGTC CGCGGCTCCG GGGTTGAAAC TCACTCCGGG GCTGTGGACG
28201 TCGGCTTACC TTCGCAAATT TGTACCTGAG GACTACCACG CCCACGAGAT TAGGTTCTAC
28261 GAAGACCAAT CCCGCCCGCC AAATGCCGAG CTTACCGCCT GCGTCATTAC CCAGGGCCAC
28321 ATCCTTGGCC AATTGCAAGC CATCAACAAA GCCCGCCAAG AGTTTCTGCT ACGAAAGGGA
28381 CGGGGGGTTT ACCTGGACCC CCAGTCCGGC GAGGAGCTCA ACCCAATCCC CCCGCCGCCG
28441 CAGCCCTATC AGCAGCCGCG GCGCCTTGCT TCCCAGGATG GCACCCAAAA AGAAGCTGCA
28501 GCTGCCGCCG CCGCGACCCA CGGACGAGGA GGAATACTGG GACAGTCAGG CAGAGGAGGT
28561 TTTGGACGAG GAGGAGGAGA TGATGGAAGA CTGGGACAGC CTAGACGAAG CTTCCGAGGC
28621 CGAAGAGGTG TCAGACGAAA CACCGTCACC CTCGGTCGCA TTCCCCTCGC CGGCGCCCCA
28681 GAAATTGGCA ACCGTTCCCA GCATCGCTAC AACCTCCGCT CCTCAGGCGC CGCCGGCACT
28741 GCCTGTTCGC CGACCCAACC GTAGATGGGA CACCACTGGA ACCAGGGCCG GTAAGTCTAA
28801 GCAGCCGCCG CCGTTAGCCC AAGAGCAACA ACAGCGCCAA GGCTACCGCT CGTGGCGCGG
28861 GCACAAGAAC GCCATAGTTG CTTGCTTGCA AGACTGTGGG GGCAACATCT CCTTGGCCCG
28921 CCGCTTTCTT CTCTACCATC ACGGCGTGGC CTTCCCCCGT AACATCCTGC ATTACTACCG
28981 TCATCTCTAC AGCCCCTACT GCACCGGCGG CAGCGGCAGC GGCAGCAACA GCAGCGGTCA
29041 CACAGAAGCA AAGGCGACCG GATAGCAAGA CTCTGACAAA GCCCAAGAAA TCCACAGCGG
29101 CGGCAGCAGC AGGAGGAGGA GCGCTGCGTC TGGCGCCCAA CGAACCCGTA TCGACCCGCG
29161 AGCTTAGAAA TAGGATTTTT CCCACTCTGT ATGCTATATT TCAACAAAGC AGGGGCCAAG

Nucleotide Sequence Analysis (cont.)

```
29221 AACAAGAGCT GAAAATAAAA AACAGGTCTC TGCGCTCCCT CACCCGCAGC TGCCTGTATC
29281 ACAAAAGCGA AGATCAGCTT CGGCGCAGGC TGGAAGACGC GGAGGCTCTC TTCAGCAAAT
29341 ACTGCGCGCT GACTCTTAAG GACTAGTTTC GCGCCCTTTC TCAAATTTAA GCGCGAAAAC
29401 TACGTCATCT CCAGCGGCCA CACCCGGCGC CAGCACCTGT CGTCAGCGCC ATTATGAGCA
29461 AGGAAATTCC CACGCCCTAC ATGTGGAGTT ACCAGCCACA AATGGGACTT GCGGCTGGAG
29521 CTGCCCAAGA CTACTCAACC CGAATAAACT ACATGAGCGC GGGACCCCAC ATGATATCCC
29581 GGGTCAACGG AATCCGCGCC CACCGAAACC GAATTCTCCT CGAACAGGCG GCTATTACCA
29641 CCACACCTCG TAATAACCTT AATCCCCGTA GTTGCCCCGC TGCCCTGGTG TACCAGGAAA
29701 GTCCCGCTCC CACCACTGTG GTACTTCCCA GAGACGCCCA GGCCGAAGTT CAGATGACTA
29761 ACTCAGGGGC GCAGCTTGCG GGCGGCTTTC GTCACAGGGT GCGGTCGCCC GGGCAGGGTA
29821 TAACTCACCT GAAAATCAGA GGGCGAGGTA TTCAGCTCAA CGACGAGTCG GTGAGCTCCT
29881 CTCTTGGTCT CCGTCCGGAC GGGACATTTC AGATCGGCGG CGCTGGCCGC TCTTCATTTA
29941 CGCCCCGTCA GGCGATCCTA ACTCTGCAGA CCTCGTCCTC GGAGCCGCGC TCCGGAGGCA
30001 TTGGAACTCT ACAATTTATT GAGGAGTTCG TGCCTTCGGT TTACTTCAAC CCCTTTTCTG
30061 GACCTCCCGG CCACTACCCG GACCAGTTTA TTCCCAACTT TGACGCGGTG AAAGACTCGG
30121 CGGACGGCTA CGACTGAATG ACCAGTGGAG AGGCAGAGCG ACTGCGCCTG ACACACCTCG
30181 ACCACTGCCG CCGCCACAAG TGCTTTGCCC GCGGCTCCGG TGAGTTTTGT TACTTTGAAT
30241 TGCCCGAAGA GCATATCGAG GGCCCGGCGC ACGGCGTCCG GCTCACCACC CAGGTAGAGC
30301 TTACACGTAG CCTGATTCGG GAGTTTACCA AGCGCCCCCT GCTAGTGGAG CGGGAGCGGC
30361 GTCCCTGTGT TCTGACCGTG GTTTGCAACT GTCCTAACCC TGGATTACAT CAAGATCTTT
30421 GTTGTCATCT CTGTGCTGAG TATAATAAAT ACAGAAATTA GAATCTACTG GGGCTCCTGT
30481 CGCCATCCTG TGAACGCCAC CGTTTTTACC CACCCAAAGC AGACCAAAGC AAACCTCACC
30541 TCCGGTTTGC ACAAGCGGGC CAATAAGTAC CTTACCTGGT ACTTTAACGG CTCTTCATTT
30601 GTAATTTACA ACAGTTTCCA GCGAGACGAA GTAAGTTTGC CACACAACCT TCTCGGCTTC
30661 AACTACACCG TCAAGAAAAA CACCACCACC ACCACCCTCC TCACCTGCCG GGAACGTACG
30721 AGTGCGTCAC CGGTTGCTGC GCCCACACCT ACAGCCTGAG CGTAACCAGA CATTACTCCC
30781 ATTTTTCCAA AACAGGAGGT GAGCTCAACT CCCGGAACTC AGGTCAAAAA AGCATTTTGC
30841 GGGGTGCTGG GATTTTTTAA TTAAGTATAT GAGCAATTCA AGTAACTCTA CAAGCTTGTC
30901 TAATTTTTCT GGAATTGGGG TCGGGGTTAT CCTTACTCTT GTAATTCTGT TTATTCTTAT
30961 ACTAGCACTT CTGTGCCTTA GGGTTGCCGC CTGCTGCACG CACGTTTGTA CCTATTGTCA
31021 GCTTTTTAAA CGCTGGGGGC AACATCCAAG ATGAGGTACA TGATTTTAGG CTTGCTCGCC
31081 CTTGCGGCAG TCTGCAGCGC TGCCAAAAAG GTTGAGTTTA AGGAACCAGC TTGCAATGTT
31141 ACATTTAAAT CAGAAGCTAA TGAATGCACT ACTCTTATAA AATGCACCAC AGAACATGAA
31201 AAGCTTATTA TTCGCCACAA AGACAAAATT GGCAAGTATG CTGTATATGC TATTTGGCAG
31261 CCAGGTGACA CTAACGACTA TAATGTCACA GTCTTCCAAG GTGAAAATCG TAAAACTTTT
31321 ATGTATAAAT TTCCATTTTA TGAAATGTGC GATATTACCA TGTACATGAG CAAACAGTAC
31381 AAGTTGTGGC CCCCACAAAA GTGTTTAGAG AACACTGGCA CCTTTTGTTC CACCGCTCTG
31441 CTTATTACAG CGCTTGCTTT GGTATGTACC TTACTTTATC TCAAATACAA AAGCAGACGC
31501 AGTTTTATTG ATGAAAAGAA AATGCCTTGA TTTTCCGCTT GCTTGTATTC CCCTGGACAA
31561 TTTACTCTAT GTGGCATATG CTCCAGGCGG GCAAGATTAT ACCCCACAACC TTCAAATCAA
31621 ACTTTCCTGG ACGTTAGCGC CTGATTTCTG CCAGCGCCTG CACTGCAAAT TTGATCAAAC
31681 CCAGCTTCAG CTTGCCTGCT CCAGAGATGA CCGGCTCAAC CATCGCGCCC ACAACGGACT
31741 ATCGCAACAC CACTGCTACC GGACTAACAT CTGCCCTAAA TTTACCCCAA GTTCATGCCT
31801 TTGTCAATGA CTGGGCGAGC TTGGACATGT GGTGGTTTTC CATAGCGCTT ATGTTTGTTT
31861 GCCTTATTAT TATGTGGCTT ATTTGTTGCC TAAAGCGCAG ACGCGCCAGA CCCCCCATCT
31921 ATAGGCCTAT CATTGTGCTC AACCCACACA ATGAAAAAAT TCATAGATTG GACGGTCTGA
31981 AACCATGTTC TCTTCTTTTA CAGTATGATT AAATGAGACA TGATTCCTCG AGTTCTTATA
32041 TTATTGACCC TTGTTGCGCT TTTCTGTGCG TGCTCTACAT TGGCCGCGGT CGCTCACATC
32101 GAAGTAGATT GCATCCCACC TTTCACAGTT TACCTGCTTT ACGGATTTGT CACCCTTATC
32161 CTCATCTGCA GCCTCGTCAC TGTAGTCATC GCCTTCATTC AGTTCATTGA CTGGGTTTGT
32221 GTGCGCATTG CGTACCTCAG GCACCATCCG CAATACAGAG ACAGGACTAT AGCTGATCTT
32281 CTCAGAATTC TTTAATTATG AAACGGAGTG TCATTTTTGT TTTGCTGATT TTTTGCGCCC
32341 TACCTGTGCT TTGCTCCCAA ACCTCAGCGC GTCCCAAAAG ACATATTTCC TGCAGATTCA
32401 CTCAAATATG GAACATTCCC AGCTGCTACA ACAAACAGAG CGATTTGTCA GAAGCCTGGT
32461 TATACGCCAT CATCTCTGTC ATGGTTTTTT GCAGTACCAT TTTTGCCCTA GCCATATATC
32521 CATACCTTGA CATTGGCTGG AATGCCATAG ATGCCATGAA CCACCCTACT TTCCCAGTGC
32581 CCGCTGTCAT ACCACTGCAA CAGGTTATTG CCCCAATCAA TCAGCCTCGC CCCCCTTCTC
```

Nucleotide Sequence Analysis (cont.)

32641 CCACCCCCAC TGAGATTAGC TACTTTAATT TGACAGGTGG AGATGACTGA ATCTCTAGAT
32701 CTAGAATTGG ATGGAATTAA CACCGAACAG CGCCTACTAG AAAGGCCCAA GGCGGCCGTCC
32761 GAGCGAGAAC GCCTAAAACA AGAAGTTGAA GACATGGTTA ACCTACACCA GTGTAAAAGA
32821 GGTATCTTTT GTGTGGTCAA GCAGGCCAAA CTTACCTACG AAAAAACCAC TACCGGCAAC
32881 CGCCTCAGCT ACAAGCTACC CACCCAGCGC CAAAAACTGG TGCTTATGGT GGGAGAAAAA
32941 CCTATCACCG TCACCCAGCA CTGGCAGAA ACAGAGGGCT GCCTGCACTT CCCCTATCAG
33001 GGTCCAGAGG ACCTCTGCAC TCTTATTAAA ACCATGTGTG GTATTAGAGA TCTTATTCCA
33061 TTCAACTAAC ATAAACACAC AATAAATTAC TTACTTAAAA TCAGTCAGCA AATCTTTGTC
33121 CAGCTTATTC AGCATCACCT CCTTTCCTTC CTCCCAACTC TGGTATCTCA GCCGCCTTTT
33181 AGCTGCAAAC TTTCTCCAAA GTTTAAATGG GATGTCAAAT TCCTCATGTT CTTGTCCCTC
33241 CGCACCCACT ATCTTCATAT TGTTGCAGAT GAAACGCGCC AGACCGTCTG AAGACACCTT
33301 CAACCCCGTG TATCCATATG ACACAGAAAC CGGGCCTCCA ACTGTGCCCT TTCTTACCCC
33361 TCCATTTGTT TCACCCAATG GTTTCCAAGA AAGTCCCCCT GGAGTTCTCT CTCTACGCGT
33421 CTCCGAACCT TTGGACACCT CCCACGGCAT GCTTGCGCTT AAAATGGGCA GCGGTCTTAC
33481 CCTAGACAAG GCCGGAAACC TCACCTCCCA AAATGTAACC ACTGTTACTC AGCCACTTAA
33541 AAAAACAAAG TCAAACATAA GTTTGGACAC CTCCGCACCA CTTACAATTA CCTCAGGCGC
33601 CCTAACAGTG GCAACCACCG CTCCTCTGAT AGTTACTAGC GGCGCTCTTA GCGTACAGTC
33661 ACAAGCCCCA CTGACCGTGC AAGACTCCAA ACTAAGCATT GCTACTAAAG GGCCCATTAC
33721 AGTGTCAGAT GGAAAGCTAG CCCTGCAAAC ATCAGCCCCC CTCTCTGGCA GTGACAGCGA
33781 CACCCTTACT GTAACTGCAT CACCCCCGCT AACTACTGCC ACGGGTAGCT TGGGCATTAA
33841 CATGGAAGAT CCTATTTATG TAAATAATGG AAAAATAGGA ATTAAAATAA GCGGTCCTTT
33901 GCAAGTAGCA CAAAACTCCG ATACACTAAC AGTAGTTACT GGACCAGGTG TCACCGTTGA
33961 ACAAAACTCC CTTAGAACCA AAGTTGCAGG AGCTATTGGT TATGATTCAT CAAACAACAT
34021 GGAAATTAAA ACGGGCGGTG GCATGCGTAT AAATAACAAC TTGTTAATTC TAGATGTGGA
34081 TTACCCATTT GATGCTCAAA CAAAACTACG TCTTAAACTG GGGCAGGGAC GCCTGTATAT
34141 TAATGCATCT CATAACTTGG ACATAAACTA TAACAGAGGC CTATACCTTT TTAATGCATC
34201 AAACAATACT AAAAAACTGG AAGTTAGCAT AAAAAAATCC AGTGGACTAA ACTTTGATAA
34261 TACTGCCATA GCTATAAATG CAGGAAAGGG TCTGGAGTTT GATACAAACA CATCTGAGTC
34321 TCCAGATATC AACCCAATAA AAACTAAAAT TGGCTCTGGC ATTGATTACA ATGAAAACGG
34381 TGCCATGATT ACTAAACTTG GAGCGGGTTT AAGCTTTGAC AACTCAGGGG CCATTACAAT
34441 AGGAAACAAA AATGATGACA AACTTACCCT GTGGACAACC CCAGACCCAT CTCCTAACTG
34501 CAGAATTCAT TCAGATAATG ACTGCAAATT TACTTTGGTT CTTACAAAAT GTGGGAGTCA
34561 AGTACTAGCT ACTGTAGCTG CTTTGGCTGT ATCTGGAGAT CTTTCATCCA TGACAGGCAC
34621 CGTTGCAAGT GTTAGTATAT TCCTTAGATT TGACCAAAAC GGTGTTCTAA TGGAGAACTC
34681 CTCACTTAAA AAACATTACT GGAACTTTAG AAATGGGAAC TCAACTAATG CAAATCCATA
34741 CACAAATGCA GTTGGATTTA TGCCTAACCT TCTAGCCTAT CCAAAAACCC AAAGTCAAAC
34801 TGCTAAAAAT AACATTGTCA GTCAAGTTTA CTTGCATGGT GATAAAACTA AACCTATGAT
34861 ACTTACCATT ACACTTAATG GCACTAGTGA ATCCACAGAA ACTAGCGAGG TAAGCACTTA
34921 CTCTATGTCT TTTACATGGT CCTGGGAAAG TGGAAAATAC ACCACTGAAA CTTTTGCTAC
34981 CAACTCTTAC ACCTTCTCCT ACATTGCCCA GGAATAAAGA ATCGTGAACC TGTTGCATGT
35041 TATGTTTCAA CGTGGGATCC TTTATTATAG GGGAAGTCCA CGCCTACATG GGGGTAGAGT
35101 CATAATCGTG CATCAGGATA GGGCGGTGGT GCTGCAGCAG CGCGCGAATA AACTGCTGCC
35161 GCCGCCGCTC CGTCCTGCAG GAATACAACA TGGCAGTGGT CTCCTCAGCG ATGATTCGCA
35221 CCGCCCGCAG CATGAGACGC CTTGTCCTCC GGGCACAGCA GCGCACCCTG ATCTCACTTA
35281 AATCAGCACA GTAACTGCAG CACAGCACCA CAATATTGTT CAAAATCCCA CAGTGCAAGG
35341 CGCTGTATCC AAAGCTCATG GCGGGGACCA CAGAACCCAC GTGGCCATCA TACCACAAGC
35401 GCAGGTAGAT TAAGTGGCGA CCCCTCATAA ACACGCTGGA CATAAACATT ACCTCTTTTG
35461 GCATGTTGTA ATTCACCACC TCCCGGTACC ATATAAACCT CTGATTAAAC ATGGCGCCAT
35521 CCACCACCAT CCTAAACCAG CTGGCCAAAA CCTGCCCGCC GGCTATGCAC TGCAGGGAAC
35581 CGGGACTGGA ACAATGACAG TGGAGAGCCC AGGACTCGTA ACCATGGATC ATCATGCTCG
35641 TCATGATATC AATGTTGGCA CAACACAGGC ACACGTGCAT ACACTTCCTC AGGATTACAA
35701 GCTCCTCCCG CGTCAGAACC ATATCCCAGG GAACAACCCA TTCCTGAATC AGCGTAAATC
35761 CCACACTGCA GGGAAGACCT CGCACGTAAC TCACGTTGTG CATTGTCAAA GTGTTACATT
35821 CGGGCAGCAG CGGATGATCC TCCAGTATGG TAGCGCGGGT CTCTGTCTCA AAAGGAGGTA
35881 GGCGATCCCT ACTGTACGGA GTGCGCCGAG ACAACCGAGA TCGTGTTGGT CGTAGTGTCA
35941 TGCCAAATGG AACGCCGGAG GTAGTCATAT TTCATCGACA CGGCACCAGC TCAATCAGTC
36001 ACAGTGTAAA AAGGGCCAAG TACAGAGCGA GTATATATAG GACTAAAAAA TGACGTAACG

Nucleotide Sequence Analysis (cont.)

```
36061 GTTAAAGTCC ACAAAAAACA CCCAGAAAAC CGCACGCGAA CCTACGCCCA GAAACGAAAG
36121 CCAAAAAACC CACAACTTCC TCAAATCTTC ACTTCCGTTT TCCCACGATA CGTCACTTCC
36181 CATTTTAAAA AAACTACAAT TCCCAATACA TGCAAGTTAC TCCGCCCTAA AACCTACGTC
36241 ACCCGCCCCG TTCCCACGCC CCGCGCCACG TCACAAACTC CACCCCCTCA TTATCATATT
36301 GGCTTCAATC CAAAATAAGG TATATTATGA TGATG
//
```

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS: Gregory, R.J., Armentano, D., Couture, L.A., Smith, A.E.

(ii) TITLE OF INVENTION: GENE THERAPY FOR CYSTIC FIBROSIS

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: LAHIVE & COCKFIELD
(B) STREET: 60 STATE STREET, SUITE 510
(C) CITY: BOSTON
(D) STATE: MASSACHUSETTS
(E) COUNTRY: USA
(F) ZIP: 02109

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: ASCII

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 02-DEC-1993
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/985,478
(B) FILING DATE: 02-DEC-1992
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Hanley, Elizabeth A.
(B) REGISTRATION NUMBER: 33,505
(C) REFERENCE/DOCKET NUMBER: NZI-014CP2PC

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 227-7400
(B) TELEFAX: (617) 227-5941

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6129 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 133..4572

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AATTGGAAGC AAATGACATC ACAGCAGGTC AGAGAAAAAG GGTTGAGCGG CAGGCACCCA      60

GAGTAGTAGG TCTTTGGCAT TAGGAGCTTG AGCCCAGACG GCCCTAGCAG GGACCCCAGC     120

GCCCGAGAGA CC ATG CAG AGG TCG CCT CTG GAA AAG GCC AGC GTT GTC           168
              Met Gln Arg Ser Pro Leu Glu Lys Ala Ser Val Val
                1               5                  10

TCC AAA CTT TTT TTC AGC TGG ACC AGA CCA ATT TTG AGG AAA GGA TAC        216
Ser Lys Leu Phe Phe Ser Trp Thr Arg Pro Ile Leu Arg Lys Gly Tyr
          15                  20                  25

AGA CAG CGC CTG GAA TTG TCA GAC ATA TAC CAA ATC CCT TCT GTT GAT        264
Arg Gln Arg Leu Glu Leu Ser Asp Ile Tyr Gln Ile Pro Ser Val Asp
      30                  35                  40

TCT GCT GAC AAT CTA TCT GAA AAA TTG GAA AGA GAA TGG GAT AGA GAG        312
Ser Ala Asp Asn Leu Ser Glu Lys Leu Glu Arg Glu Trp Asp Arg Glu
  45                  50                  55                  60

CTG GCT TCA AAG AAA AAT CCT AAA CTC ATT AAT GCC CTT CGG CGA TGT        360
Leu Ala Ser Lys Lys Asn Pro Lys Leu Ile Asn Ala Leu Arg Arg Cys
                  65                  70                  75

TTT TTC TGG AGA TTT ATG TTC TAT GGA ATC TTT TTA TAT TTA GGG GAA        408
Phe Phe Trp Arg Phe Met Phe Tyr Gly Ile Phe Leu Tyr Leu Gly Glu
              80                  85                  90

GTC ACC AAA GCA GTA CAG CCT CTC TTA CTG GGA AGA ATC ATA GCT TCC        456
Val Thr Lys Ala Val Gln Pro Leu Leu Leu Gly Arg Ile Ile Ala Ser
          95                  100                 105

TAT GAC CCG GAT AAC AAG GAG GAA CGC TCT ATC GCG ATT TAT CTA GGC        504
Tyr Asp Pro Asp Asn Lys Glu Glu Arg Ser Ile Ala Ile Tyr Leu Gly
      110                 115                 120

ATA GGC TTA TGC CTT CTC TTT ATT GTG AGG ACA CTG CTC CTA CAC CCA        552
Ile Gly Leu Cys Leu Leu Phe Ile Val Arg Thr Leu Leu Leu His Pro
125                 130                 135                 140

GCC ATT TTT GGC CTT CAT CAC ATT GGA ATG CAG ATG AGA ATA GCT ATG        600
Ala Ile Phe Gly Leu His His Ile Gly Met Gln Met Arg Ile Ala Met
                  145                 150                 155

TTT AGT TTG ATT TAT AAG AAG ACT TTA AAG CTG TCA AGC CGT GTT CTA        648
Phe Ser Leu Ile Tyr Lys Lys Thr Leu Lys Leu Ser Ser Arg Val Leu
              160                 165                 170
```

```
GAT AAA ATA AGT ATT GGA CAA CTT GTT AGT CTC CTT TCC AAC AAC CTG      696
Asp Lys Ile Ser Ile Gly Gln Leu Val Ser Leu Leu Ser Asn Asn Leu
        175                 180                 185

AAC AAA TTT GAT GAA GGA CTT GCA TTG GCA CAT TTC GTG TGG ATC GCT      744
Asn Lys Phe Asp Glu Gly Leu Ala Leu Ala His Phe Val Trp Ile Ala
    190                 195                 200

CCT TTG CAA GTG GCA CTC CTC ATG GGG CTA ATC TGG GAG TTG TTA CAG      792
Pro Leu Gln Val Ala Leu Leu Met Gly Leu Ile Trp Glu Leu Leu Gln
205                 210                 215                 220

GCG TCT GCC TTC TGT GGA CTT GGT TTC CTG ATA GTC CTT GCC CTT TTT      840
Ala Ser Ala Phe Cys Gly Leu Gly Phe Leu Ile Val Leu Ala Leu Phe
                225                 230                 235

CAG GCT GGG CTA GGG AGA ATG ATG ATG AAG TAC AGA GAT CAG AGA GCT      888
Gln Ala Gly Leu Gly Arg Met Met Met Lys Tyr Arg Asp Gln Arg Ala
            240                 245                 250

GGG AAG ATC AGT GAA AGA CTT GTG ATT ACC TCA GAA ATG ATT GAA AAT      936
Gly Lys Ile Ser Glu Arg Leu Val Ile Thr Ser Glu Met Ile Glu Asn
        255                 260                 265

ATC CAA TCT GTT AAG GCA TAC TGC TGG GAA GAA GCA ATG GAA AAA ATG      984
Ile Gln Ser Val Lys Ala Tyr Cys Trp Glu Glu Ala Met Glu Lys Met
    270                 275                 280

ATT GAA AAC TTA AGA CAA ACA GAA CTG AAA CTG ACT CGG AAG GCA GCC     1032
Ile Glu Asn Leu Arg Gln Thr Glu Leu Lys Leu Thr Arg Lys Ala Ala
285                 290                 295                 300

TAT GTG AGA TAC TTC AAT AGC TCA GCC TTC TTC TTC TCA GGG TTC TTT     1080
Tyr Val Arg Tyr Phe Asn Ser Ser Ala Phe Phe Phe Ser Gly Phe Phe
                305                 310                 315

GTG GTG TTT TTA TCT GTG CTT CCC TAT GCA CTA ATC AAA GGA ATC ATC     1128
Val Val Phe Leu Ser Val Leu Pro Tyr Ala Leu Ile Lys Gly Ile Ile
            320                 325                 330

CTC CGG AAA ATA TTC ACC ACC ATC TCA TTC TGC ATT GTT CTG CGC ATG     1176
Leu Arg Lys Ile Phe Thr Thr Ile Ser Phe Cys Ile Val Leu Arg Met
        335                 340                 345

GCG GTC ACT CGG CAA TTT CCC TGG GCT GTA CAA ACA TGG TAT GAC TCT     1224
Ala Val Thr Arg Gln Phe Pro Trp Ala Val Gln Thr Trp Tyr Asp Ser
    350                 355                 360

CTT GGA GCA ATA AAC AAA ATA CAG GAT TTC TTA CAA AAG CAA GAA TAT     1272
Leu Gly Ala Ile Asn Lys Ile Gln Asp Phe Leu Gln Lys Gln Glu Tyr
365                 370                 375                 380

AAG ACA TTG GAA TAT AAC TTA ACG ACT ACA GAA GTA GTG ATG GAG AAT     1320
Lys Thr Leu Glu Tyr Asn Leu Thr Thr Thr Glu Val Val Met Glu Asn
                385                 390                 395
```

```
GTA ACA GCC TTC TGG GAG GAG GGA TTT GGG GAA TTA TTT GAG AAA GCA      1368
Val Thr Ala Phe Trp Glu Glu Gly Phe Gly Glu Leu Phe Glu Lys Ala
            400                 405                 410

AAA CAA AAC AAT AAC AAT AGA AAA ACT TCT AAT GGT GAT GAC AGC CTC      1416
Lys Gln Asn Asn Asn Asn Arg Lys Thr Ser Asn Gly Asp Asp Ser Leu
        415                 420                 425

TTC TTC AGT AAT TTC TCA CTT CTT GGT ACT CCT GTC CTG AAA GAT ATT      1464
Phe Phe Ser Asn Phe Ser Leu Leu Gly Thr Pro Val Leu Lys Asp Ile
    430                 435                 440

AAT TTC AAG ATA GAA AGA GGA CAG TTG TTG GCG GTT GCT GGA TCC ACT      1512
Asn Phe Lys Ile Glu Arg Gly Gln Leu Leu Ala Val Ala Gly Ser Thr
445                 450                 455                 460

GGA GCA GGC AAG ACT TCA CTT CTA ATG ATG ATT ATG GGA GAA CTG GAG      1560
Gly Ala Gly Lys Thr Ser Leu Leu Met Met Ile Met Gly Glu Leu Glu
                465                 470                 475

CCT TCA GAG GGT AAA ATT AAG CAC AGT GGA AGA ATT TCA TTC TGT TCT      1608
Pro Ser Glu Gly Lys Ile Lys His Ser Gly Arg Ile Ser Phe Cys Ser
            480                 485                 490

CAG TTT TCC TGG ATT ATG CCT GGC ACC ATT AAA GAA AAT ATC ATC TTT      1656
Gln Phe Ser Trp Ile Met Pro Gly Thr Ile Lys Glu Asn Ile Ile Phe
        495                 500                 505

GGT GTT TCC TAT GAT GAA TAT AGA TAC AGA AGC GTC ATC AAA GCA TGC      1704
Gly Val Ser Tyr Asp Glu Tyr Arg Tyr Arg Ser Val Ile Lys Ala Cys
    510                 515                 520

CAA CTA GAA GAG GAC ATC TCC AAG TTT GCA GAG AAA GAC AAT ATA GTT      1752
Gln Leu Glu Glu Asp Ile Ser Lys Phe Ala Glu Lys Asp Asn Ile Val
525                 530                 535                 540

CTT GGA GAA GGT GGA ATC ACA CTG AGT GGA GGT CAA CGA GCA AGA ATT      1800
Leu Gly Glu Gly Gly Ile Thr Leu Ser Gly Gly Gln Arg Ala Arg Ile
                545                 550                 555

TCT TTA GCA AGA GCA GTA TAC AAA GAT GCT GAT TTG TAT TTA TTA GAC      1848
Ser Leu Ala Arg Ala Val Tyr Lys Asp Ala Asp Leu Tyr Leu Leu Asp
            560                 565                 570

TCT CCT TTT GGA TAC CTA GAT GTT TTA ACA GAA AAA GAA ATA TTT GAA      1896
Ser Pro Phe Gly Tyr Leu Asp Val Leu Thr Glu Lys Glu Ile Phe Glu
        575                 580                 585

AGC TGT GTC TGT AAA CTG ATG GCT AAC AAA ACT AGG ATT TTG GTC ACT      1944
Ser Cys Val Cys Lys Leu Met Ala Asn Lys Thr Arg Ile Leu Val Thr
    590                 595                 600

TCT AAA ATG GAA CAT TTA AAG AAA GCT GAC AAA ATA TTA ATT TTG CAT      1992
Ser Lys Met Glu His Leu Lys Lys Ala Asp Lys Ile Leu Ile Leu His
605                 610                 615                 620
```

```
GAA GGT AGC AGC TAT TTT TAT GGG ACA TTT TCA GAA CTC CAA AAT CTA      2040
Glu Gly Ser Ser Tyr Phe Tyr Gly Thr Phe Ser Glu Leu Gln Asn Leu
                625                 630                 635

CAG CCA GAC TTT AGC TCA AAA CTC ATG GGA TGT GAT TCT TTC GAC CAA      2088
Gln Pro Asp Phe Ser Ser Lys Leu Met Gly Cys Asp Ser Phe Asp Gln
            640                 645                 650

TTT AGT GCA GAA AGA AGA AAT TCA ATC CTA ACT GAG ACC TTA CAC CGT      2136
Phe Ser Ala Glu Arg Arg Asn Ser Ile Leu Thr Glu Thr Leu His Arg
        655                 660                 665

TTC TCA TTA GAA GGA GAT GCT CCT GTC TCC TGG ACA GAA ACA AAA AAA      2184
Phe Ser Leu Glu Gly Asp Ala Pro Val Ser Trp Thr Glu Thr Lys Lys
    670                 675                 680

CAA TCT TTT AAA CAG ACT GGA GAG TTT GGG GAA AAA AGG AAG AAT TCT      2232
Gln Ser Phe Lys Gln Thr Gly Glu Phe Gly Glu Lys Arg Lys Asn Ser
685                 690                 695                 700

ATT CTC AAT CCA ATC AAC TCT ATA CGA AAA TTT TCC ATT GTG CAA AAG      2280
Ile Leu Asn Pro Ile Asn Ser Ile Arg Lys Phe Ser Ile Val Gln Lys
                705                 710                 715

ACT CCC TTA CAA ATG AAT GGC ATC GAA GAG GAT TCT GAT GAG CCT TTA      2328
Thr Pro Leu Gln Met Asn Gly Ile Glu Glu Asp Ser Asp Glu Pro Leu
            720                 725                 730

GAG AGA AGG CTG TCC TTA GTA CCA GAT TCT GAG CAG GGA GAG GCG ATA      2376
Glu Arg Arg Leu Ser Leu Val Pro Asp Ser Glu Gln Gly Glu Ala Ile
        735                 740                 745

CTG CCT CGC ATC AGC GTG ATC AGC ACT GGC CCC ACG CTT CAG GCA CGA      2424
Leu Pro Arg Ile Ser Val Ile Ser Thr Gly Pro Thr Leu Gln Ala Arg
    750                 755                 760

AGG AGG CAG TCT GTC CTG AAC CTG ATG ACA CAC TCA GTT AAC CAA GGT      2472
Arg Arg Gln Ser Val Leu Asn Leu Met Thr His Ser Val Asn Gln Gly
765                 770                 775                 780

CAG AAC ATT CAC CGA AAG ACA ACA GCA TCC ACA CGA AAA GTG TCA CTG      2520
Gln Asn Ile His Arg Lys Thr Thr Ala Ser Thr Arg Lys Val Ser Leu
                785                 790                 795

GCC CCT CAG GCA AAC TTG ACT GAA CTG GAT ATA TAT TCA AGA AGG TTA      2568
Ala Pro Gln Ala Asn Leu Thr Glu Leu Asp Ile Tyr Ser Arg Arg Leu
            800                 805                 810

TCT CAA GAA ACT GGC TTG GAA ATA AGT GAA GAA ATT AAC GAA GAA GAC      2616
Ser Gln Glu Thr Gly Leu Glu Ile Ser Glu Glu Ile Asn Glu Glu Asp
        815                 820                 825

TTA AAG GAG TGC CTT TTT GAT GAT ATG GAG AGC ATA CCA GCA GTG ACT      2664
Leu Lys Glu Cys Leu Phe Asp Asp Met Glu Ser Ile Pro Ala Val Thr
    830                 835                 840
```

```
ACA TGG AAC ACA TAC CTT CGA TAT ATT ACT GTC CAC AAG AGC TTA ATT      2712
Thr Trp Asn Thr Tyr Leu Arg Tyr Ile Thr Val His Lys Ser Leu Ile
845                 850                 855                 860

TTT GTG CTA ATT TGG TGC TTA GTA ATT TTT CTG GCA GAG GTG GCT GCT      2760
Phe Val Leu Ile Trp Cys Leu Val Ile Phe Leu Ala Glu Val Ala Ala
                865                 870                 875

TCT TTG GTT GTG CTG TGG CTC CTT GGA AAC ACT CCT CTT CAA GAC AAA      2808
Ser Leu Val Val Leu Trp Leu Leu Gly Asn Thr Pro Leu Gln Asp Lys
            880                 885                 890

GGG AAT AGT ACT CAT AGT AGA AAT AAC AGC TAT GCA GTG ATT ATC ACC      2856
Gly Asn Ser Thr His Ser Arg Asn Asn Ser Tyr Ala Val Ile Ile Thr
        895                 900                 905

AGC ACC AGT TCG TAT TAT GTG TTT TAC ATT TAC GTG GGA GTA GCC GAC      2904
Ser Thr Ser Ser Tyr Tyr Val Phe Tyr Ile Tyr Val Gly Val Ala Asp
    910                 915                 920

ACT TTG CTT GCT ATG GGA TTC TTC AGA GGT CTA CCA CTG GTG CAT ACT      2952
Thr Leu Leu Ala Met Gly Phe Phe Arg Gly Leu Pro Leu Val His Thr
925                 930                 935                 940

CTA ATC ACA GTG TCG AAA ATT TTA CAC CAC AAA ATG TTA CAT TCT GTT      3000
Leu Ile Thr Val Ser Lys Ile Leu His His Lys Met Leu His Ser Val
                945                 950                 955

CTT CAA GCA CCT ATG TCA ACC CTC AAC ACG TTG AAA GCA GGT GGG ATT      3048
Leu Gln Ala Pro Met Ser Thr Leu Asn Thr Leu Lys Ala Gly Gly Ile
            960                 965                 970

CTT AAT AGA TTC TCC AAA GAT ATA GCA ATT TTG GAT GAC CTT CTG CCT      3096
Leu Asn Arg Phe Ser Lys Asp Ile Ala Ile Leu Asp Asp Leu Leu Pro
        975                 980                 985

CTT ACC ATA TTT GAC TTC ATC CAG TTG TTA TTA ATT GTG ATT GGA GCT      3144
Leu Thr Ile Phe Asp Phe Ile Gln Leu Leu Leu Ile Val Ile Gly Ala
    990                 995                 1000

ATA GCA GTT GTC GCA GTT TTA CAA CCC TAC ATC TTT GTT GCA ACA GTG      3192
Ile Ala Val Val Ala Val Leu Gln Pro Tyr Ile Phe Val Ala Thr Val
1005                1010                1015                1020

CCA GTG ATA GTG GCT TTT ATT ATG TTG AGA GCA TAT TTC CTC CAA ACC      3240
Pro Val Ile Val Ala Phe Ile Met Leu Arg Ala Tyr Phe Leu Gln Thr
                1025                1030                1035

TCA CAG CAA CTC AAA CAA CTG GAA TCT GAA GGC AGG AGT CCA ATT TTC      3288
Ser Gln Gln Leu Lys Gln Leu Glu Ser Glu Gly Arg Ser Pro Ile Phe
            1040                1045                1050

ACT CAT CTT GTT ACA AGC TTA AAA GGA CTA TGG ACA CTT CGT GCC TTC      3336
Thr His Leu Val Thr Ser Leu Lys Gly Leu Trp Thr Leu Arg Ala Phe
        1055                1060                1065
```

```
GGA CGG CAG CCT TAC TTT GAA ACT CTG TTC CAC AAA GCT CTG AAT TTA      3384
Gly Arg Gln Pro Tyr Phe Glu Thr Leu Phe His Lys Ala Leu Asn Leu
    1070                1075                1080

CAT ACT GCC AAC TGG TTC TTG TAC CTG TCA ACA CTG CGC TGG TTC CAA      3432
His Thr Ala Asn Trp Phe Leu Tyr Leu Ser Thr Leu Arg Trp Phe Gln
1085                1090                1095                1100

ATG AGA ATA GAA ATG ATT TTT GTC ATC TTC TTC ATT GCT GTT ACC TTC      3480
Met Arg Ile Glu Met Ile Phe Val Ile Phe Phe Ile Ala Val Thr Phe
                1105                1110                1115

ATT TCC ATT TTA ACA ACA GGA GAA GGA GAA GGA AGA GTT GGT ATT ATC      3528
Ile Ser Ile Leu Thr Thr Gly Glu Gly Glu Gly Arg Val Gly Ile Ile
            1120                1125                1130

CTG ACT TTA GCC ATG AAT ATC ATG AGT ACA TTG CAG TGG GCT GTA AAC      3576
Leu Thr Leu Ala Met Asn Ile Met Ser Thr Leu Gln Trp Ala Val Asn
        1135                1140                1145

TCC AGC ATA GAT GTG GAT AGC TTG ATG CGA TCT GTG AGC CGA GTC TTT      3624
Ser Ser Ile Asp Val Asp Ser Leu Met Arg Ser Val Ser Arg Val Phe
    1150                1155                1160

AAG TTC ATT GAC ATG CCA ACA GAA GGT AAA CCT ACC AAG TCA ACC AAA      3672
Lys Phe Ile Asp Met Pro Thr Glu Gly Lys Pro Thr Lys Ser Thr Lys
1165                1170                1175                1180

CCA TAC AAG AAT GGC CAA CTC TCG AAA GTT ATG ATT ATT GAG AAT TCA      3720
Pro Tyr Lys Asn Gly Gln Leu Ser Lys Val Met Ile Ile Glu Asn Ser
                1185                1190                1195

CAC GTG AAG AAA GAT GAC ATC TGG CCC TCA GGG GGC CAA ATG ACT GTC      3768
His Val Lys Lys Asp Asp Ile Trp Pro Ser Gly Gly Gln Met Thr Val
            1200                1205                1210

AAA GAT CTC ACA GCA AAA TAC ACA GAA GGT GGA AAT GCC ATA TTA GAG      3816
Lys Asp Leu Thr Ala Lys Tyr Thr Glu Gly Gly Asn Ala Ile Leu Glu
        1215                1220                1225

AAC ATT TCC TTC TCA ATA AGT CCT GGC CAG AGG GTG GGC CTC TTG GGA      3864
Asn Ile Ser Phe Ser Ile Ser Pro Gly Gln Arg Val Gly Leu Leu Gly
    1230                1235                1240

AGA ACT GGA TCA GGG AAG AGT ACT TTG TTA TCA GCT TTT TTG AGA CTA      3912
Arg Thr Gly Ser Gly Lys Ser Thr Leu Leu Ser Ala Phe Leu Arg Leu
1245                1250                1255                1260

CTG AAC ACT GAA GGA GAA ATC CAG ATC GAT GGT GTG TCT TGG GAT TCA      3960
Leu Asn Thr Glu Gly Glu Ile Gln Ile Asp Gly Val Ser Trp Asp Ser
                1265                1270                1275

ATA ACT TTG CAA CAG TGG AGG AAA GCC TTT GGA GTG ATA CCA CAG AAA      4008
Ile Thr Leu Gln Gln Trp Arg Lys Ala Phe Gly Val Ile Pro Gln Lys
            1280                1285                1290
```

```
GTA TTT ATT TTT TCT GGA ACA TTT AGA AAA AAC TTG GAT CCC TAT GAA      4056
Val Phe Ile Phe Ser Gly Thr Phe Arg Lys Asn Leu Asp Pro Tyr Glu
        1295                1300                1305

CAG TGG AGT GAT CAA GAA ATA TGG AAA GTT GCA GAT GAG GTT GGG CTC      4104
Gln Trp Ser Asp Gln Glu Ile Trp Lys Val Ala Asp Glu Val Gly Leu
    1310                1315                1320

AGA TCT GTG ATA GAA CAG TTT CCT GGG AAG CTT GAC TTT GTC CTT GTG      4152
Arg Ser Val Ile Glu Gln Phe Pro Gly Lys Leu Asp Phe Val Leu Val
1325                1330                1335                1340

GAT GGG GGC TGT GTC CTA AGC CAT GGC CAC AAG CAG TTG ATG TGC TTG      4200
Asp Gly Gly Cys Val Leu Ser His Gly His Lys Gln Leu Met Cys Leu
                1345                1350                1355

GCT AGA TCT GTT CTC AGT AAG GCG AAG ATC TTG CTG CTT GAT GAA CCC      4248
Ala Arg Ser Val Leu Ser Lys Ala Lys Ile Leu Leu Leu Asp Glu Pro
            1360                1365                1370

AGT GCT CAT TTG GAT CCA GTA ACA TAC CAA ATA ATT AGA AGA ACT CTA      4296
Ser Ala His Leu Asp Pro Val Thr Tyr Gln Ile Ile Arg Arg Thr Leu
        1375                1380                1385

AAA CAA GCA TTT GCT GAT TGC ACA GTA ATT CTC TGT GAA CAC AGG ATA      4344
Lys Gln Ala Phe Ala Asp Cys Thr Val Ile Leu Cys Glu His Arg Ile
    1390                1395                1400

GAA GCA ATG CTG GAA TGC CAA CAA TTT TTG GTC ATA GAA GAG AAC AAA      4392
Glu Ala Met Leu Glu Cys Gln Gln Phe Leu Val Ile Glu Glu Asn Lys
1405                1410                1415                1420

GTG CGG CAG TAC GAT TCC ATC CAG AAA CTG CTG AAC GAG AGG AGC CTC      4440
Val Arg Gln Tyr Asp Ser Ile Gln Lys Leu Leu Asn Glu Arg Ser Leu
                1425                1430                1435

TTC CGG CAA GCC ATC AGC CCC TCC GAC AGG GTG AAG CTC TTT CCC CAC      4488
Phe Arg Gln Ala Ile Ser Pro Ser Asp Arg Val Lys Leu Phe Pro His
            1440                1445                1450

CGG AAC TCA AGC AAG TGC AAG TCT AAG CCC CAG ATT GCT GCT CTG AAA      4536
Arg Asn Ser Ser Lys Cys Lys Ser Lys Pro Gln Ile Ala Ala Leu Lys
        1455                1460                1465

GAG GAG ACA GAA GAA GAG GTG CAA GAT ACA AGG CTT TAGAGAGCAG          4582
Glu Glu Thr Glu Glu Glu Val Gln Asp Thr Arg Leu
    1470                1475                1480

CATAAATGTT GACATGGGAC ATTTGCTCAT GGAATTGGAG CTCGTGGGAC AGTCACCTCA   4642

TGGAATTGGA GCTCGTGGAA CAGTTACCTC TGCCTCAGAA AACAAGGATG AATTAAGTTT   4702

TTTTTTAAAA AAGAAACATT TGGTAAGGGG AATTGAGGAC ACTGATATGG GTCTTGATAA   4762

ATGGCTTCCT GGCAATAGTC AAATTGTGTG AAAGGTACTT CAAATCCTTG AAGATTTACC   4822

ACTTGTGTTT TGCAAGCCAG ATTTTCCTGA AAACCCTTGC CATGTGCTAG TAATTGGAGA   4882
```

GGCAGCTCTA AATGTCAATC AGCCTAGTTG ATCAGCTTAT TGTCTAGTGA AACTCGTTAA 4942

TTTGTAGTGT TGGAGAAGAA CTGAAATCAT ACTTCTTAGG GTTATGATTA AGTAATGATA 5002

ACTGGAAACT TCAGCGGTTT ATATAAGCTT GTATTCCTTT TTCTCTCCTC TCCCCATGAT 5062

GTTTAGAAAC ACAACTATAT TGTTTGCTAA GCATTCCAAC TATCTCATTT CCAAGCAAGT 5122

ATTAGAATAC CACAGGAACC ACAAGACTGC ACATCAAAAT ATGCCCCATT CAACATCTAG 5182

TGAGCAGTCA GGAAAGAGAA CTTCCAGATC CTGGAAATCA GGGTTAGTAT TGTCCAGGTC 5242

TACCAAAAAT CTCAATATTT CAGATAATCA CAATACATCC CTTACCTGGG AAAGGGCTGT 5302

TATAATCTTT CACAGGGGAC AGGATGGTTC CCTTGATGAA GAAGTTGATA TGCCTTTTCC 5362

CAACTCCAGA AAGTGACAAG CTCACAGACC TTTGAACTAG AGTTTAGCTG GAAAAGTATG 5422

TTAGTGCAAA TTGTCACAGG ACAGCCCTTC TTTCCACAGA AGCTCCAGGT AGAGGGTGTG 5482

TAAGTAGATA GGCCATGGGC ACTGTGGGTA GACACACATG AAGTCCAAGC ATTTAGATGT 5542

ATAGGTTGAT GGTGGTATGT TTTCAGGCTA GATGTATGTA CTTCATGCTG TCTACACTAA 5602

GAGAGAATGA GAGACACACT GAAGAAGCAC CAATCATGAA TTAGTTTTAT ATGCTTCTGT 5662

TTTATAATTT TGTGAAGCAA AATTTTTTCT CTAGGAAATA TTTATTTTAA TAATGTTTCA 5722

AACATATATT ACAATGCTGT ATTTTAAAAG AATGATTATG AATTACATTT GTATAAAATA 5782

ATTTTTATAT TTGAAATATT GACTTTTTAT GGCACTAGTA TTTTTATGAA ATATTATGTT 5842

AAAACTGGGA CAGGGGAGAA CCTAGGGTGA TATTAACCAG GGGCCATGAA TCACCTTTTG 5902

GTCTGGAGGG AAGCCTTGGG GCTGATCGAG TTGTTGCCCA CAGCTGTATG ATTCCCAGCC 5962

AGACACAGCC TCTTAGATGC AGTTCTGAAG AAGATGGTAC CACCAGTCTG ACTGTTTCCA 6022

TCAAGGGTAC ACTGCCTTCT CAACTCCAAA CTGACTCTTA AGAAGACTGC ATTATATTTA 6082

TTACTGTAAG AAAATATCAC TTGTCAATAA AATCCATACA TTTGTGT 6129

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1480 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gln Arg Ser Pro Leu Glu Lys Ala Ser Val Val Ser Lys Leu Phe
1 5 10 15

```
Phe Ser Trp Thr Arg Pro Ile Leu Arg Lys Gly Tyr Arg Gln Arg Leu
            20                  25                  30

Glu Leu Ser Asp Ile Tyr Gln Ile Pro Ser Val Asp Ser Ala Asp Asn
        35                  40                  45

Leu Ser Glu Lys Leu Glu Arg Glu Trp Asp Arg Glu Leu Ala Ser Lys
    50                  55                  60

Lys Asn Pro Lys Leu Ile Asn Ala Leu Arg Arg Cys Phe Phe Trp Arg
65                  70                  75                  80

Phe Met Phe Tyr Gly Ile Phe Leu Tyr Leu Gly Glu Val Thr Lys Ala
                85                  90                  95

Val Gln Pro Leu Leu Leu Gly Arg Ile Ile Ala Ser Tyr Asp Pro Asp
            100                 105                 110

Asn Lys Glu Glu Arg Ser Ile Ala Ile Tyr Leu Gly Ile Gly Leu Cys
        115                 120                 125

Leu Leu Phe Ile Val Arg Thr Leu Leu Leu His Pro Ala Ile Phe Gly
    130                 135                 140

Leu His His Ile Gly Met Gln Met Arg Ile Ala Met Phe Ser Leu Ile
145                 150                 155                 160

Tyr Lys Lys Thr Leu Lys Leu Ser Ser Arg Val Leu Asp Lys Ile Ser
                165                 170                 175

Ile Gly Gln Leu Val Ser Leu Leu Ser Asn Asn Leu Asn Lys Phe Asp
            180                 185                 190

Glu Gly Leu Ala Leu Ala His Phe Val Trp Ile Ala Pro Leu Gln Val
        195                 200                 205

Ala Leu Leu Met Gly Leu Ile Trp Glu Leu Leu Gln Ala Ser Ala Phe
    210                 215                 220

Cys Gly Leu Gly Phe Leu Ile Val Leu Ala Leu Phe Gln Ala Gly Leu
225                 230                 235                 240

Gly Arg Met Met Met Lys Tyr Arg Asp Gln Arg Ala Gly Lys Ile Ser
                245                 250                 255

Glu Arg Leu Val Ile Thr Ser Glu Met Ile Glu Asn Ile Gln Ser Val
            260                 265                 270

Lys Ala Tyr Cys Trp Glu Glu Ala Met Glu Lys Met Ile Glu Asn Leu
        275                 280                 285

Arg Gln Thr Glu Leu Lys Leu Thr Arg Lys Ala Ala Tyr Val Arg Tyr
    290                 295                 300

Phe Asn Ser Ser Ala Phe Phe Phe Ser Gly Phe Phe Val Val Phe Leu
305                 310                 315                 320
```

Ser Val Leu Pro Tyr Ala Leu Ile Lys Gly Ile Ile Leu Arg Lys Ile
325 330 335

Phe Thr Thr Ile Ser Phe Cys Ile Val Leu Arg Met Ala Val Thr Arg
340 345 350

Gln Phe Pro Trp Ala Val Gln Thr Trp Tyr Asp Ser Leu Gly Ala Ile
355 360 365

Asn Lys Ile Gln Asp Phe Leu Gln Lys Gln Glu Tyr Lys Thr Leu Glu
370 375 380

Tyr Asn Leu Thr Thr Thr Glu Val Val Met Glu Asn Val Thr Ala Phe
385 390 395 400

Trp Glu Glu Gly Phe Gly Glu Leu Phe Glu Lys Ala Lys Gln Asn Asn
405 410 415

Asn Asn Arg Lys Thr Ser Asn Gly Asp Asp Ser Leu Phe Phe Ser Asn
420 425 430

Phe Ser Leu Leu Gly Thr Pro Val Leu Lys Asp Ile Asn Phe Lys Ile
435 440 445

Glu Arg Gly Gln Leu Leu Ala Val Ala Gly Ser Thr Gly Ala Gly Lys
450 455 460

Thr Ser Leu Leu Met Met Ile Met Gly Glu Leu Glu Pro Ser Glu Gly
465 470 475 480

Lys Ile Lys His Ser Gly Arg Ile Ser Phe Cys Ser Gln Phe Ser Trp
485 490 495

Ile Met Pro Gly Thr Ile Lys Glu Asn Ile Ile Phe Gly Val Ser Tyr
500 505 510

Asp Glu Tyr Arg Tyr Arg Ser Val Ile Lys Ala Cys Gln Leu Glu Glu
515 520 525

Asp Ile Ser Lys Phe Ala Glu Lys Asp Asn Ile Val Leu Gly Glu Gly
530 535 540

Gly Ile Thr Leu Ser Gly Gly Gln Arg Ala Arg Ile Ser Leu Ala Arg
545 550 555 560

Ala Val Tyr Lys Asp Ala Asp Leu Tyr Leu Leu Asp Ser Pro Phe Gly
565 570 575

Tyr Leu Asp Val Leu Thr Glu Lys Glu Ile Phe Glu Ser Cys Val Cys
580 585 590

Lys Leu Met Ala Asn Lys Thr Arg Ile Leu Val Thr Ser Lys Met Glu
595 600 605

His Leu Lys Lys Ala Asp Lys Ile Leu Ile Leu His Glu Gly Ser Ser
610 615 620

```
Tyr Phe Tyr Gly Thr Phe Ser Glu Leu Gln Asn Leu Gln Pro Asp Phe
625                 630                 635                 640

Ser Ser Lys Leu Met Gly Cys Asp Ser Phe Asp Gln Phe Ser Ala Glu
                645                 650                 655

Arg Arg Asn Ser Ile Leu Thr Glu Thr Leu His Arg Phe Ser Leu Glu
            660                 665                 670

Gly Asp Ala Pro Val Ser Trp Thr Glu Thr Lys Lys Gln Ser Phe Lys
        675                 680                 685

Gln Thr Gly Glu Phe Gly Glu Lys Arg Lys Asn Ser Ile Leu Asn Pro
    690                 695                 700

Ile Asn Ser Ile Arg Lys Phe Ser Ile Val Gln Lys Thr Pro Leu Gln
705                 710                 715                 720

Met Asn Gly Ile Glu Glu Asp Ser Asp Glu Pro Leu Glu Arg Arg Leu
                725                 730                 735

Ser Leu Val Pro Asp Ser Glu Gln Gly Glu Ala Ile Leu Pro Arg Ile
            740                 745                 750

Ser Val Ile Ser Thr Gly Pro Thr Leu Gln Ala Arg Arg Arg Gln Ser
        755                 760                 765

Val Leu Asn Leu Met Thr His Ser Val Asn Gln Gly Gln Asn Ile His
    770                 775                 780

Arg Lys Thr Thr Ala Ser Thr Arg Lys Val Ser Leu Ala Pro Gln Ala
785                 790                 795                 800

Asn Leu Thr Glu Leu Asp Ile Tyr Ser Arg Arg Leu Ser Gln Glu Thr
                805                 810                 815

Gly Leu Glu Ile Ser Glu Glu Ile Asn Glu Glu Asp Leu Lys Glu Cys
            820                 825                 830

Leu Phe Asp Asp Met Glu Ser Ile Pro Ala Val Thr Thr Trp Asn Thr
        835                 840                 845

Tyr Leu Arg Tyr Ile Thr Val His Lys Ser Leu Ile Phe Val Leu Ile
    850                 855                 860

Trp Cys Leu Val Ile Phe Leu Ala Glu Val Ala Ala Ser Leu Val Val
865                 870                 875                 880

Leu Trp Leu Leu Gly Asn Thr Pro Leu Gln Asp Lys Gly Asn Ser Thr
                885                 890                 895

His Ser Arg Asn Asn Ser Tyr Ala Val Ile Ile Thr Ser Thr Ser Ser
            900                 905                 910

Tyr Tyr Val Phe Tyr Ile Tyr Val Gly Val Ala Asp Thr Leu Leu Ala
        915                 920                 925
```

Met Gly Phe Phe Arg Gly Leu Pro Leu Val His Thr Leu Ile Thr Val
930 935 940

Ser Lys Ile Leu His His Lys Met Leu His Ser Val Leu Gln Ala Pro
945 950 955 960

Met Ser Thr Leu Asn Thr Leu Lys Ala Gly Gly Ile Leu Asn Arg Phe
965 970 975

Ser Lys Asp Ile Ala Ile Leu Asp Asp Leu Leu Pro Leu Thr Ile Phe
980 985 990

Asp Phe Ile Gln Leu Leu Leu Ile Val Ile Gly Ala Ile Ala Val Val
995 1000 1005

Ala Val Leu Gln Pro Tyr Ile Phe Val Ala Thr Val Pro Val Ile Val
1010 1015 1020

Ala Phe Ile Met Leu Arg Ala Tyr Phe Leu Gln Thr Ser Gln Gln Leu
1025 1030 1035 1040

Lys Gln Leu Glu Ser Glu Gly Arg Ser Pro Ile Phe Thr His Leu Val
1045 1050 1055

Thr Ser Leu Lys Gly Leu Trp Thr Leu Arg Ala Phe Gly Arg Gln Pro
1060 1065 1070

Tyr Phe Glu Thr Leu Phe His Lys Ala Leu Asn Leu His Thr Ala Asn
1075 1080 1085

Trp Phe Leu Tyr Leu Ser Thr Leu Arg Trp Phe Gln Met Arg Ile Glu
1090 1095 1100

Met Ile Phe Val Ile Phe Phe Ile Ala Val Thr Phe Ile Ser Ile Leu
1105 1110 1115 1120

Thr Thr Gly Glu Gly Glu Gly Arg Val Gly Ile Ile Leu Thr Leu Ala
1125 1130 1135

Met Asn Ile Met Ser Thr Leu Gln Trp Ala Val Asn Ser Ser Ile Asp
1140 1145 1150

Val Asp Ser Leu Met Arg Ser Val Ser Arg Val Phe Lys Phe Ile Asp
1155 1160 1165

Met Pro Thr Glu Gly Lys Pro Thr Lys Ser Thr Lys Pro Tyr Lys Asn
1170 1175 1180

Gly Gln Leu Ser Lys Val Met Ile Ile Glu Asn Ser His Val Lys Lys
1185 1190 1195 1200

Asp Asp Ile Trp Pro Ser Gly Gly Gln Met Thr Val Lys Asp Leu Thr
1205 1210 1215

Ala Lys Tyr Thr Glu Gly Gly Asn Ala Ile Leu Glu Asn Ile Ser Phe
1220 1225 1230

Ser Ile Ser Pro Gly Gln Arg Val Gly Leu Leu Gly Arg Thr Gly Ser
1235 1240 1245

Gly Lys Ser Thr Leu Leu Ser Ala Phe Leu Arg Leu Leu Asn Thr Glu
1250 1255 1260

Gly Glu Ile Gln Ile Asp Gly Val Ser Trp Asp Ser Ile Thr Leu Gln
1265 1270 1275 1280

Gln Trp Arg Lys Ala Phe Gly Val Ile Pro Gln Lys Val Phe Ile Phe
1285 1290 1295

Ser Gly Thr Phe Arg Lys Asn Leu Asp Pro Tyr Glu Gln Trp Ser Asp
1300 1305 1310

Gln Glu Ile Trp Lys Val Ala Asp Glu Val Gly Leu Arg Ser Val Ile
1315 1320 1325

Glu Gln Phe Pro Gly Lys Leu Asp Phe Val Leu Val Asp Gly Gly Cys
1330 1335 1340

Val Leu Ser His Gly His Lys Gln Leu Met Cys Leu Ala Arg Ser Val
1345 1350 1355 1360

Leu Ser Lys Ala Lys Ile Leu Leu Leu Asp Glu Pro Ser Ala His Leu
1365 1370 1375

Asp Pro Val Thr Tyr Gln Ile Ile Arg Arg Thr Leu Lys Gln Ala Phe
1380 1385 1390

Ala Asp Cys Thr Val Ile Leu Cys Glu His Arg Ile Glu Ala Met Leu
1395 1400 1405

Glu Cys Gln Gln Phe Leu Val Ile Glu Glu Asn Lys Val Arg Gln Tyr
1410 1415 1420

Asp Ser Ile Gln Lys Leu Leu Asn Glu Arg Ser Leu Phe Arg Gln Ala
1425 1430 1435 1440

Ile Ser Pro Ser Asp Arg Val Lys Leu Phe Pro His Arg Asn Ser Ser
1445 1450 1455

Lys Cys Lys Ser Lys Pro Gln Ile Ala Ala Leu Lys Glu Glu Thr Glu
1460 1465 1470

Glu Glu Val Gln Asp Thr Arg Leu
1475 1480

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5635 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CATCATCAAT AATATACCTT ATTTTGGATT GAAGCCAATA TGATAATGAG GGGGTGGAGT 60

TTGTGACGTG GCGCGGGGCG TGGGAACGGG GCGGGTGACG TAGTAGTGTG GCGGAAGTGT 120

GATGTTGCAA GTGTGGCGGA ACACATGTAA GCGCCGGATG TGGTAAAAGT GACGTTTTTG 180

GTGTGCGCCG GTGTATACGG GAAGTGACAA TTTTCGCGCG GTTTTAGGCG GATGTTGTAG 240

TAAATTTGGG CGTAACCAAG TAATGTTTGG CCATTTTCGC GGGAAAACTG AATAAGAGGA 300

AGTGAAATCT GAATAATTCT GTGTTACTCA TAGCGCGTAA TATTTGTCTA GGGCCGCGGG 360

GACTTTGACC GTTTACGTGG AGACTCGCCC AGGTGTTTTT CTCAGGTGTT TTCCGCGTTC 420

CGGGTCAAAG TTGGCGTTTT ATTATTATAG TCAGCTGACG CGCAGTGTAT TTATACCCGG 480

TGAGTTCCTC AAGAGGCCAC TCTTGAGTGC CAGCGAGTAG AGTTTTCTCC TCCGAGCCGC 540

TCCGAGCTAG TAACGGCCGC CAGTGTGCTG CAGATATCAA AGTCGACGGT ACCCGAGAGA 600

CCATGCAGAG GTCGCCTCTG GAAAAGGCCA GCGTTGTCTC CAAACTTTTT TTCAGCTGGA 660

CCAGACCAAT TTTGAGGAAA GGATACAGAC AGCGCCTGGA ATTGTCAGAC ATATACCAAA 720

TCCCTTCTGT TGATTCTGCT GACAATCTAT CTGAAAAATT GGAAAGAGAA TGGGATAGAG 780

AGCTGGCTTC AAAGAAAAAT CCTAAACTCA TTAATGCCCT TCGGCGATGT TTTTTCTGGA 840

GATTTATGTT CTATGGAATC TTTTTATATT TAGGGGAAGT CACCAAAGCA GTACAGCCTC 900

TCTTACTGGG AAGAATCATA GCTTCCTATG ACCCGGATAA CAAGGAGGAA CGCTCTATCG 960

CGATTTATCT AGGCATAGGC TTATGCCTTC TCTTTATTGT GAGGACACTG CTCCTACACC 1020

CAGCCATTTT TGGCCTTCAT CACATTGGAA TGCAGATGAG AATAGCTATG TTTAGTTTGA 1080

TTTATAAGAA GACTTTAAAG CTGTCAAGCC GTGTTCTAGA TAAAATAAGT ATTGGACAAC 1140

TTGTTAGTCT CCTTTCCAAC AACCTGAACA AATTTGATGA AGGACTTGCA TTGGCACATT 1200

TCGTGTGGAT CGCTCCTTTG CAAGTGGCAC TCCTCATGGG GCTAATCTGG GAGTTGTTAC 1260

AGGCGTCTGC CTTCTGTGGA CTTGGTTTCC TGATAGTCCT TGCCCTTTTT CAGGCTGGGC 1320

TAGGGAGAAT GATGATGAAG TACAGAGATC AGAGAGCTGG GAAGATCAGT GAAAGACTTG 1380

TGATTACCTC AGAAATGATT GAAAACATCC AATCTGTTAA GGCATACTGC TGGGAAGAAG 1440

CAATGGAAAA AATGATTGAA AACTTAAGAC AAACAGAACT GAAACTGACT CGGAAGGCAG 1500

CCTATGTGAG ATACTTCAAT AGCTCAGCCT TCTTCTTCTC AGGGTTCTTT GTGGTGTTTT 1560

TATCTGTGCT TCCCTATGCA CTAATCAAAG GAATCATCCT CCGGAAAATA TTCACCACCA 1620

TCTCATTCTG CATTGTTCTG CGCATGGCGG TCACTCGGCA ATTTCCCTGG GCTGTACAAA 1680

CATGGTATGA CTCTCTTGGA GCAATAAACA AAATACAGGA TTTCTTACAA AAGCAAGAAT 1740

ATAAGACATT GGAATATAAC TTAACGACTA CAGAAGTAGT GATGGAGAAT GTAACAGCCT 1800

TCTGGGAGGA GGGATTTGGG GAATTATTTG AGAAAGCAAA ACAAAACAAT AACAATAGAA 1860

AAACTTCTAA TGGTGATGAC AGCCTCTTCT TCAGTAATTT CTCACTTCTT GGTACTCCTG 1920

TCCTGAAAGA TATTAATTTC AAGATAGAAA GAGGACAGTT GTTGGCGGTT GCTGGATCCA 1980

CTGGAGCAGG CAAGACTTCA CTTCTAATGA TGATTATGGG AGAACTGGAG CCTTCAGAGG 2040

GTAAAATTAA GCACAGTGGA AGAATTTCAT TCTGTTCTCA GTTTTCCTGG ATTATGCCTG 2100

GCACCATTAA AGAAAATATC ATCTTTGGTG TTTCCTATGA TGAATATAGA TACAGAAGCG 2160

TCATCAAAGC ATGCCAACTA GAAGAGGACA TCTCCAAGTT TGCAGAGAAA GACAATATAG 2220

TTCTTGGAGA AGGTGGAATC ACACTGAGTG GAGGTCAACG AGCAAGAATT TCTTTAGCAA 2280

GAGCAGTATA CAAAGATGCT GATTTGTATT TATTAGACTC TCCTTTTGGA TACCTAGATG 2340

TTTTAACAGA AAAAGAAATA TTTGAAAGCT GTGTCTGTAA ACTGATGGCT AACAAAACTA 2400

GGATTTTGGT CACTTCTAAA ATGGAACATT TAAAGAAAGC TGACAAAATA TTAATTTTGC 2460

ATGAAGGTAG CAGCTATTTT TATGGGACAT TTTCAGAACT CCAAAATCTA CAGCCAGACT 2520

TTAGCTCAAA ACTCATGGGA TGTGATTCTT TCGACCAATT TAGTGCAGAA AGAAGAAATT 2580

CAATCCTAAC TGAGACCTTA CACCGTTTCT CATTAGAAGG AGATGCTCCT GTCTCCTGGA 2640

CAGAAACAAA AAAACAATCT TTTAAACAGA CTGGAGAGTT TGGGGAAAAA AGGAAGAATT 2700

CTATTCTCAA TCCAATCAAC TCTATACGAA AATTTTCCAT TGTGCAAAAG ACTCCCTTAC 2760

AAATGAATGG CATCGAAGAG GATTCTGATG AGCCTTTAGA GAGAAGGCTG TCCTTAGTAC 2820

CAGATTCTGA GCAGGGAGAG GCGATACTGC CTCGCATCAG CGTGATCAGC ACTGGCCCCA 2880

CGCTTCAGGC ACGAAGGAGG CAGTCTGTCC TGAACCTGAT GACACACTCA GTTAACCAAG 2940

GTCAGAACAT TCACCGAAAG ACAACAGCAT CCACACGAAA AGTGTCACTG GCCCCTCAGG 3000

CAAACTTGAC TGAACTGGAT ATATATTCAA GAAGGTTATC TCAAGAAACT GGCTTGGAAA 3060

TAAGTGAAGA AATTAACGAA GAAGACTTAA AGGAGTGCCT TTTTGATGAT ATGGAGAGCA 3120

TACCAGCAGT GACTACATGG AACACATACC TTCGATATAT TACTGTCCAC AAGAGCTTAA 3180

TTTTTGTGCT AATTTGGTGC TTAGTAATTT TTCTGGCAGA GGTGGCTGCT TCTTTGGTTG 3240

TGCTGTGGCT CCTTGGAAAC ACTCCTCTTC AAGACAAAGG GAATAGTACT CATAGTAGAA 3300

ATAACAGCTA TGCAGTGATT ATCACCAGCA CCAGTTCGTA TTATGTGTTT TACATTTACG 3360

TGGGAGTAGC CGACACTTTG CTTGCTATGG GATTCTTCAG AGGTCTACCA CTGGTGCATA 3420

CTCTAATCAC AGTGTCGAAA ATTTTACACC ACAAAATGTT ACATTCTGTT CTTCAAGCAC 3480

CTATGTCAAC CCTCAACACG TTGAAAGCAG GTGGGATTCT TAATAGATTC TCCAAAGATA 3540

TAGCAATTTT GGATGACCTT CTGCCTCTTA CCATATTTGA CTTCATCCAG TTGTTATTAA 3600

TTGTGATTGG AGCTATAGCA GTTGTCGCAG TTTTACAACC CTACATCTTT GTTGCAACAG 3660

TGCCAGTGAT AGTGGCTTTT ATTATGTTGA GAGCATATTT CCTCCAAACC TCACAGCAAC 3720

TCAAACAACT GGAATCTGAA GGCAGGAGTC CAATTTTCAC TCATCTTGTT ACAAGCTTAA 3780

AAGGACTATG GACACTTCGT GCCTTCGGAC GGCAGCCTTA CTTTGAAACT CTGTTCCACA 3840

AAGCTCTGAA TTTACATACT GCCAACTGGT TCTTGTACCT GTCAACACTG CGCTGGTTCC 3900

AAATGAGAAT AGAAATGATT TTTGTCATCT TCTTCATTGC TGTTACCTTC ATTTCCATTT 3960

TAACAACAGG AGAAGGAGAA GGAAGAGTTG GTATTATCCT GACTTTAGCC ATGAATATCA 4020

TGAGTACATT GCAGTGGGCT GTAAACTCCA GCATAGATGT GGATAGCTTG ATGCGATCTG 4080

TGAGCCGAGT CTTTAAGTTC ATTGACATGC CAACAGAAGG TAAACCTACC AAGTCAACCA 4140

AACCATACAA GAATGGCCAA CTCTCGAAAG TTATGATTAT TGAGAATTCA CACGTGAAGA 4200

AAGATGACAT CTGGCCCTCA GGGGGCCAAA TGACTGTCAA AGATCTCACA GCAAAATACA 4260

CAGAAGGTGG AAATGCCATA TTAGAGAACA TTTCCTTCTC AATAAGTCCT GGCCAGAGGG 4320

TGGGCCTCTT GGGAAGAACT GGATCAGGGA AGAGTACTTT GTTATCAGCT TTTTTGAGAC 4380

TACTGAACAC TGAAGGAGAA ATCCAGATCG ATGGTGTGTC TTGGGATTCA ATAACTTTGC 4440

AACAGTGGAG GAAAGCCTTT GGAGTGATAC CACAGAAAGT ATTTATTTTT TCTGGAACAT 4500

TTAGAAAAAA CTTGGATCCC TATGAACAGT GGAGTGATCA AGAAATATGG AAAGTTGCAG 4560

ATGAGGTTGG GCTCAGATCT GTGATAGAAC AGTTTCCTGG GAAGCTTGAC TTTGTCCTTG 4620

TGGATGGGGG CTGTGTCCTA AGCCATGGCC ACAAGCAGTT GATGTGCTTG GCTAGATCTG 4680

TTCTCAGTAA GGCGAAGATC TTGCTGCTTG ATGAACCCAG TGCTCATTTG GATCCAGTAA 4740

CATACCAAAT AATTAGAAGA ACTCTAAAAC AAGCATTTGC TGATTGCACA GTAATTCTCT 4800

GTGAACACAG GATAGAAGCA ATGCTGGAAT GCCAACAATT TTTGGTCATA GAAGAGAACA 4860

AAGTGCGGCA GTACGATTCC ATCCAGAAAC TGCTGAACGA GAGGAGCCTC TTCCGGCAAG 4920

CCATCAGCCC CTCCGACAGG GTGAAGCTCT TTCCCCACCG GAACTCAAGC AAGTGCAAGT 4980

CTAAGCCCCA GATTGCTGCT CTGAAAGAGG AGACAGAAGA AGAGGTGCAA GATACAAGGC 5040

TTTAGAGAGC AGCATAAATG TTGACATGGG ACATTTGCTC ATGGAATTGG AGGTAGCGGA 5100

TTGAGGTACT GAAATGTGTG GGCGTGGCTT AAGGGTGGGA AAGAATATAT AAGGTGGGGG 5160

TCTCATGTAG TTTTGTATCT GTTTTGCAGC AGCCGCCGCC ATGAGCGCCA ACTCGTTTGA 5220

TGGAAGCATT GTGAGCTCAT ATTTGACAAC GCGCATGCCC CCATGGGCCG GGGTGCGTCA 5280

GAATGTGATG GGCTCCAGCA TTGATGGTCG CCCCGTCCTG CCCGCAAACT CTACTACCTT 5340

GACCTACGAG ACCGTGTCTG GAACGCCGTT GGAGACTGCA GCCTCCGCCG CCGCTTCAGC 5400

CGCTGCAGCC ACCGCCCGCG GGATTGTGAC TGACTTTGCT TTCCTGAGCC CGCTTGCAAG 5460

CAGTGCAGCT TCCCGTTCAT CCGCCCGCGA TGACAAGTTG ACGGCTCTTT TGGCACAATT 5520

GGATTCTTTG ACCCGGGAAC TTAATGTCGT TTCTCAGCAG CTGTTGGATC TGCGCCAGCA 5580

GGTTTCTGCC CTGAAGGCTT CCTCCCCTCC CAATGCGGTT TAAAACATAA ATAAA 5635

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACTCTTGAGT GCCAGCGAGT AGAGTTTTCT CCTCCG 36

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCAAAGGAGC GATCCACACG AAATGTGCC 29

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTCCTCCGAG CCGCTCCGAG CTAG 24

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCAAAAATGG CTGGGTGTAG GAGCAGTGTC C 31

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGGATCCTTT ATTATAGGGG AAGTCCACGC CTAC 34

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGGGATCCAT CGATGAAATA TGACTACGTC CG 32

Claims

1. An adenovirus-based gene therapy vector comprising the genome of an adenovirus 2 serotype in which the Ela and Elb regions of the genome, which are involved in early stages of viral replication, have been deleted and replaced by genetic material of interest.

2. The adenovirus-based gene therapy vector of claim 1, wherein the genetic material of interest is DNA encoding cystic fibrosis transmembrane conductance regulator

3. The adenovirus-based gene therapy vector of claim 1 further comprising PGK promoter operably linked to the genetic material of interest.

4. The adenovirus-based gene therapy vector of claim 2 having substantially the same nucleotide sequence as shown in Table II (SEQ ID NO:3).

5. An adenovirus-based gene therapy vector comprising adenovirus inverted terminal repeat nucleotide sequences and the minimal nucleotide sequences necessary for efficient replication and packaging and genetic material of interest.

6. The adenovirus-based gene therapy vector of claim 5 having the adenovirus 2 sequences shown in Figure 17.

7. The adenovirus-based gene therapy vector of claim 5 further comprising PGK promoter operably linked to the genetic material of interest.

8. The adenovirus-based gene therapy vector of claim 5 in which the genetic material of interest is selected from the group consisting of DNA encoding: cystic fibrosis transmembrane conductance regulator, Factor VIII, and Factor IX.

9. An adenovirus-based gene therapy vector comprising an adenovirus genome which has been deleted for all E4 open reading frames, except open reading frame 6, and additionally comprising genetic material of interest.

10. The adenovirus-based gene therapy vector of claim 9 further comprising PGK promoter operably linked to the genetic material of interest.

11. The adenovirus-based gene therapy vector of claim 9 in which the Ela and Elb regions of the genome, which are involved in early stages of viral replication, have been deleted.

12. The adenovirus-based gene therapy vector of claim 9 in which the E3 region has been deleted.

13. An adenovirus-based gene therapy vector comprising an adenovirus genome which has been deleted for all E4 open reading frames, except open reading frame 3, and additionally comprising genetic material of interest.

14. The adenovirus-based gene therapy vector of claim 13 in which the Ela and Elb regions of the genome, which are involved in early stages of viral replication, have been deleted.

15. The adenovirus-based gene therapy vector of claim 13 further comprising PGK promoter operably linked to the genetic material of interest.

16. The adenovirus-based gene therapy vector of claim 13 in which the E3 region has been deleted.

17. A method for treating or preventing cystic fibrosis in a patient comprising administering to the pulmonary airways of the patient, a gene therapy vector comprising DNA encoding cystic fibrosis transmembrane conductance regulator.

18. The method of claim 17 wherein the gene therapy vector is an adenovirus-based gene therapy vector comprising the genome of an adenovirus 2 serotype in which the Ela and Elb regions of the genome, which are involved in early stages of viral replication, have been deleted and replaced by DNA encoding cystic fibrosis transmembrane conductance regulator.

19. The method of claim 17 wherein the gene therapy vector further comprises PGK promoter operably linked to the DNA encoding cystic fibrosis transmembrane conductance regulator.

20. The method of claim 17 wherein the gene therapy vector is an adenovirus-based gene therapy vector comprising adenovirus inverted terminal repeats and the minimal sequences necessary for efficient replication and packaging and DNA encoding cystic fibrosis tranmembrane conductance regulator.

21. The method of claim 20 wherein the gene therapy vector further comprises PGK promoter operably linked to the DNA encoding cystic fibrosis transmembrane conductance regulator.

22. The method of claim 17 wherein the gene therapy vector is an adenovirus-based gene therapy vector comprising an adenovirus genome which has been deleted for all E4 open reading frames, except open reading frame 6, and additionally comprising DNA encoding cystic fibrosis transmembrane conductance regulator.

23. The method of claim 22 wherein the gene therapy vector further comprises PGK promoter operably linked to the DNA encoding cystic fibrosis transmembrane conductance regulator.

24. The method of claim 17 wherein the gene therapy vector is an adenovirus-based gene therapy vector comprising an adenovirus genome which has been deleted for all E4 open reading frames, except open reading frame 6, and has been deleted for the Ela and Elb regions of the genome, which are involved in early stages of viral replication, and additionally comprising DNA encoding cystic fibrosis tranmembrane conductance regulator.

25. The method of claim 24 wherein the gene therapy vector further comprises PGK promoter operably linked to the DNA encoding cystic fibrosis transmembrane conductance regulator.

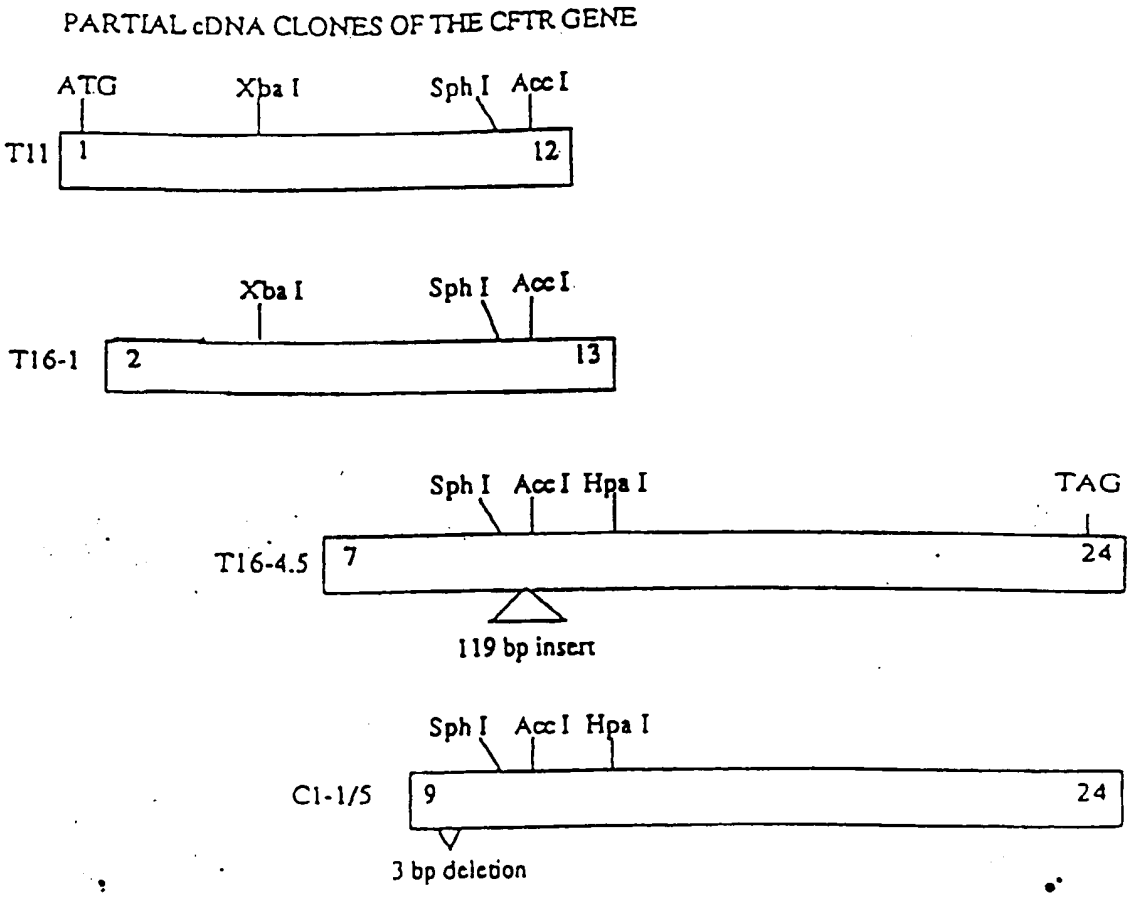

PARTIAL cDNA CLONES OF THE CFTR GENE
ATG
Xba I
Sph I
Acc I
T11
1
12
Xba I
Sph I
Acc I
T16-1
2
13
Sph I
Acc I
Hpa I
TAG
T16-4.5
7
24
119 bp insert
Sph I
Acc I
Hpa I
C1-1/5
9
24
3 bp deletion

Figure 1

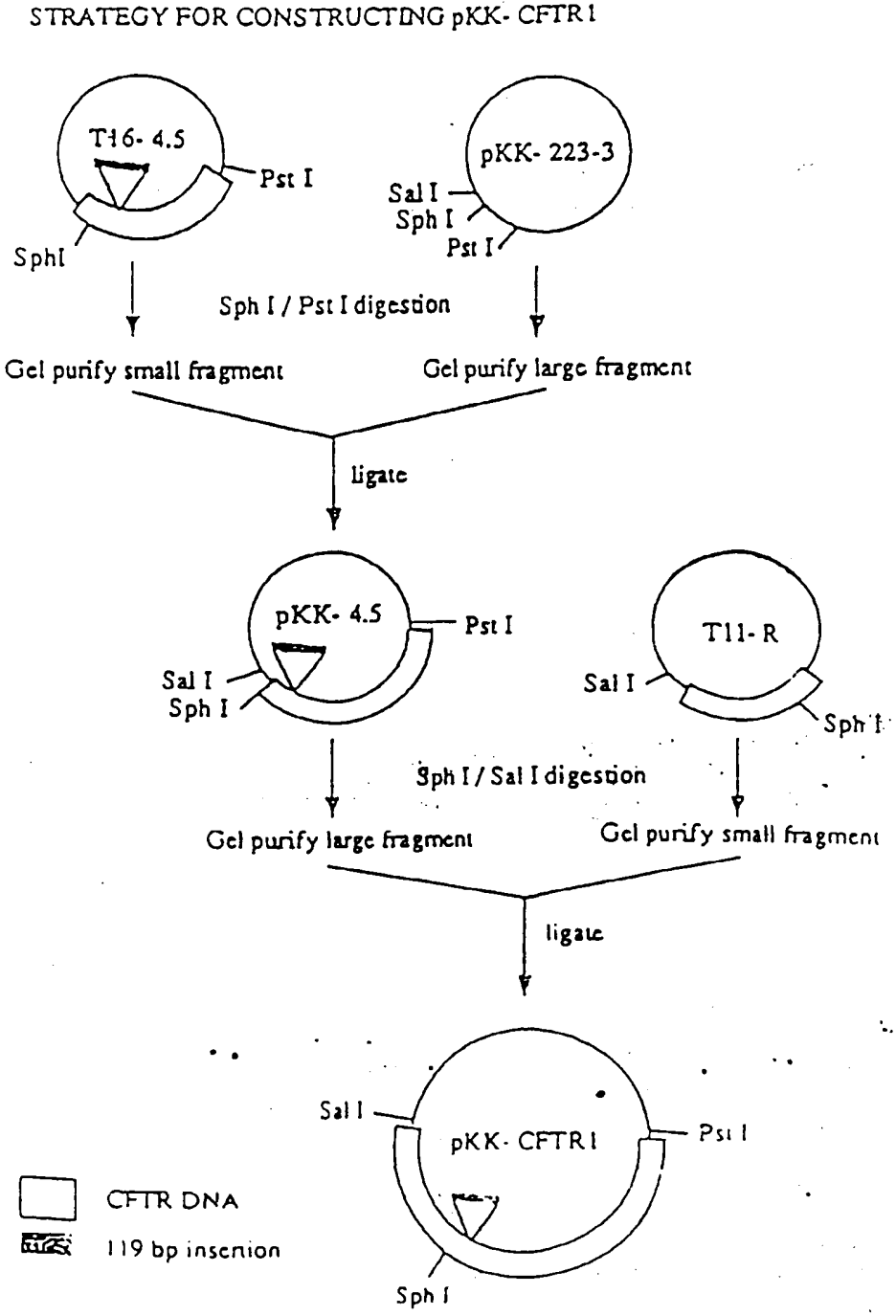

STRATEGY FOR CONSTRUCTING pKK- CFTR1
T16- 4.5
Pst I
SphI
pKK- 223-3
Sal I
Sph I
Pst I
Sph I / Pst I digestion
Gel purify small fragment
Gel purify large fragment
ligate
pKK- 4.5
Pst I
Sal I
Sph I
T11- R
Sal I
Sph I
Sph I / Sal I digestion
Gel purify large fragment
Gel purify small fragment
ligate
Sal I
pKK- CFTR1
Pst I
Sph I
CFTR DNA
119 bp insertion

Figure 2

CONSTRUCTION OF THE pKK- CFTR2 PLASMID

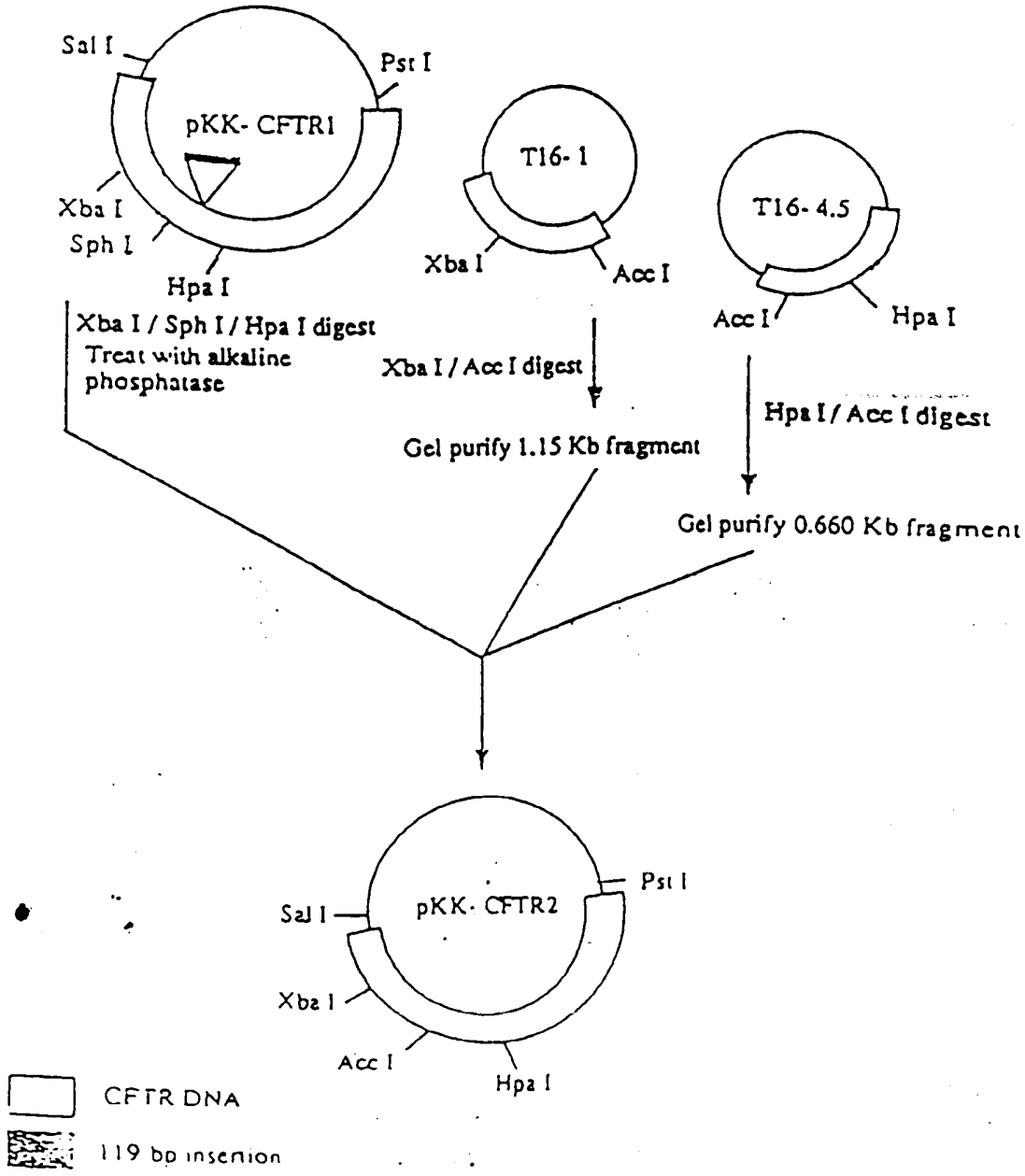


Figure 3

STRATEGY FOR CONSTRUCTING THE pSC- CFTR2 PLASMID

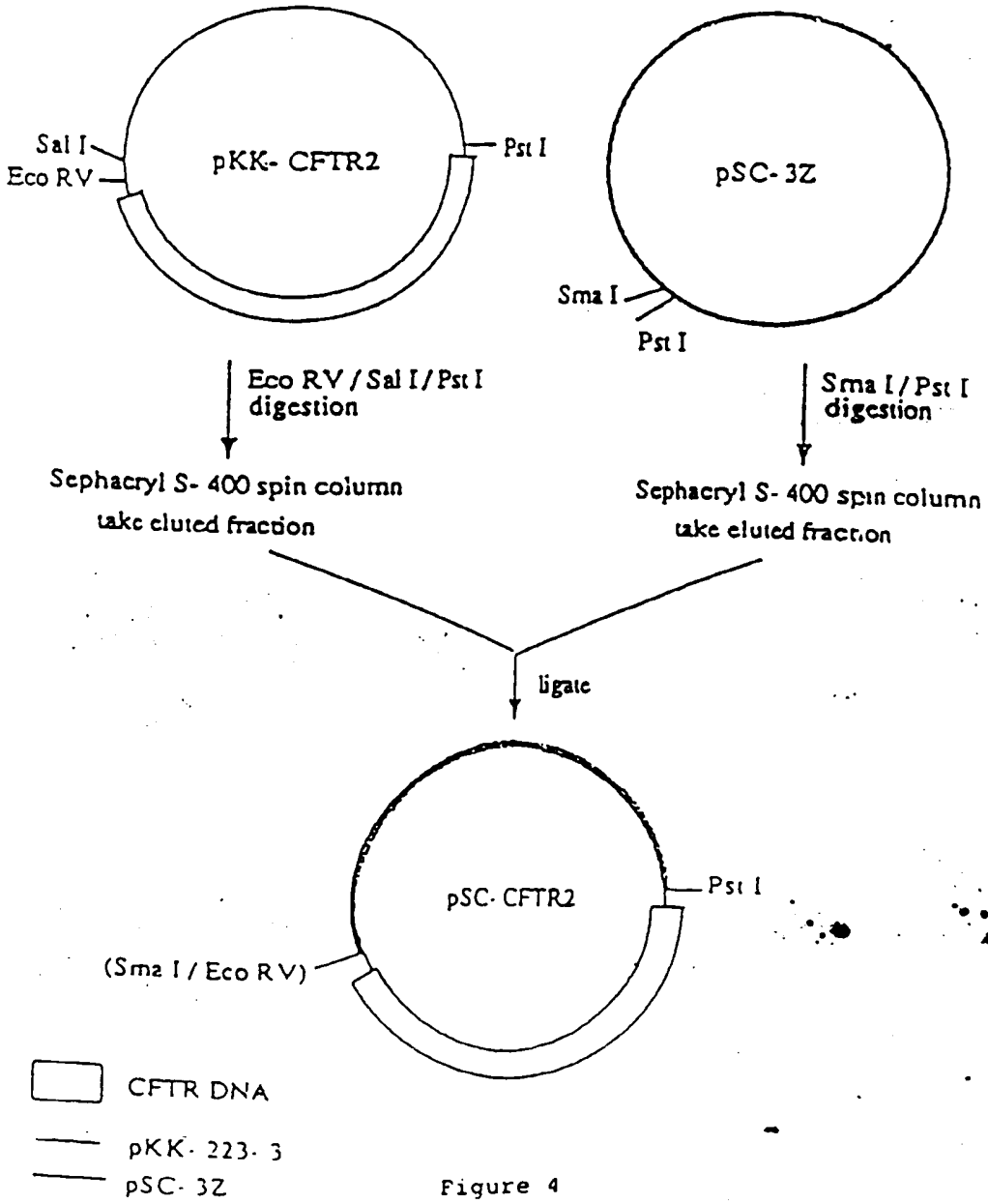


Figure 4

MAP OF pSC- CFTR2

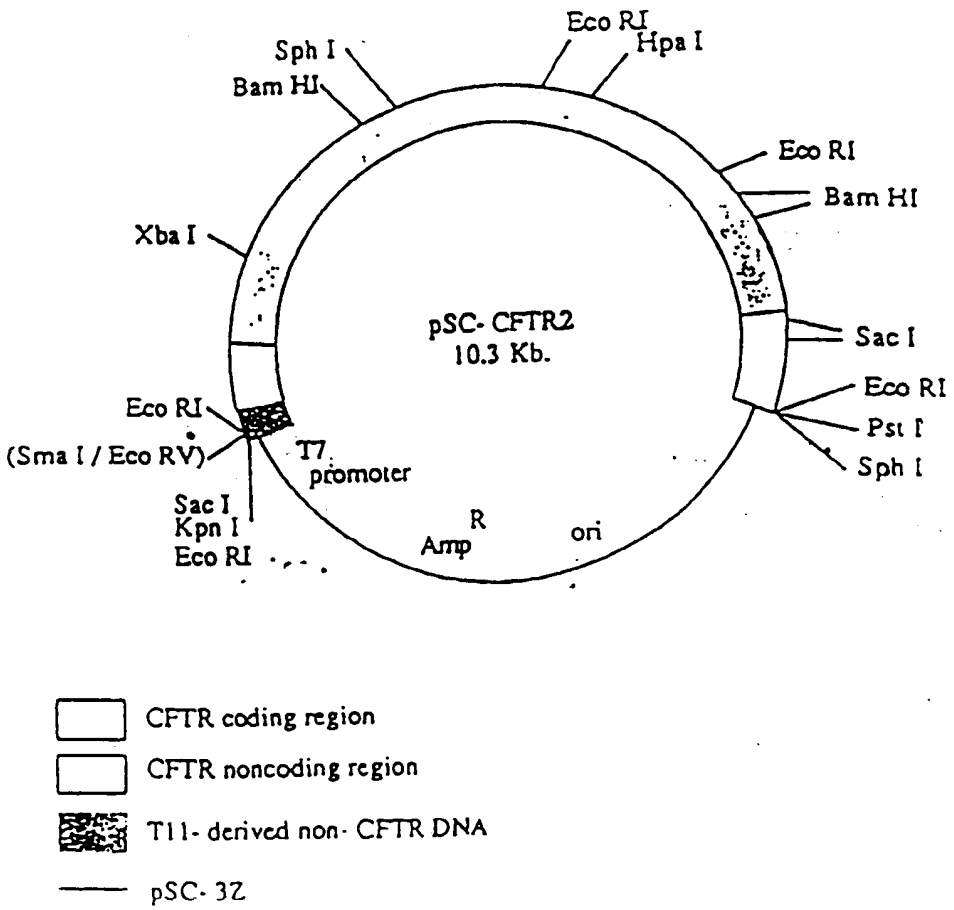


Figure 5

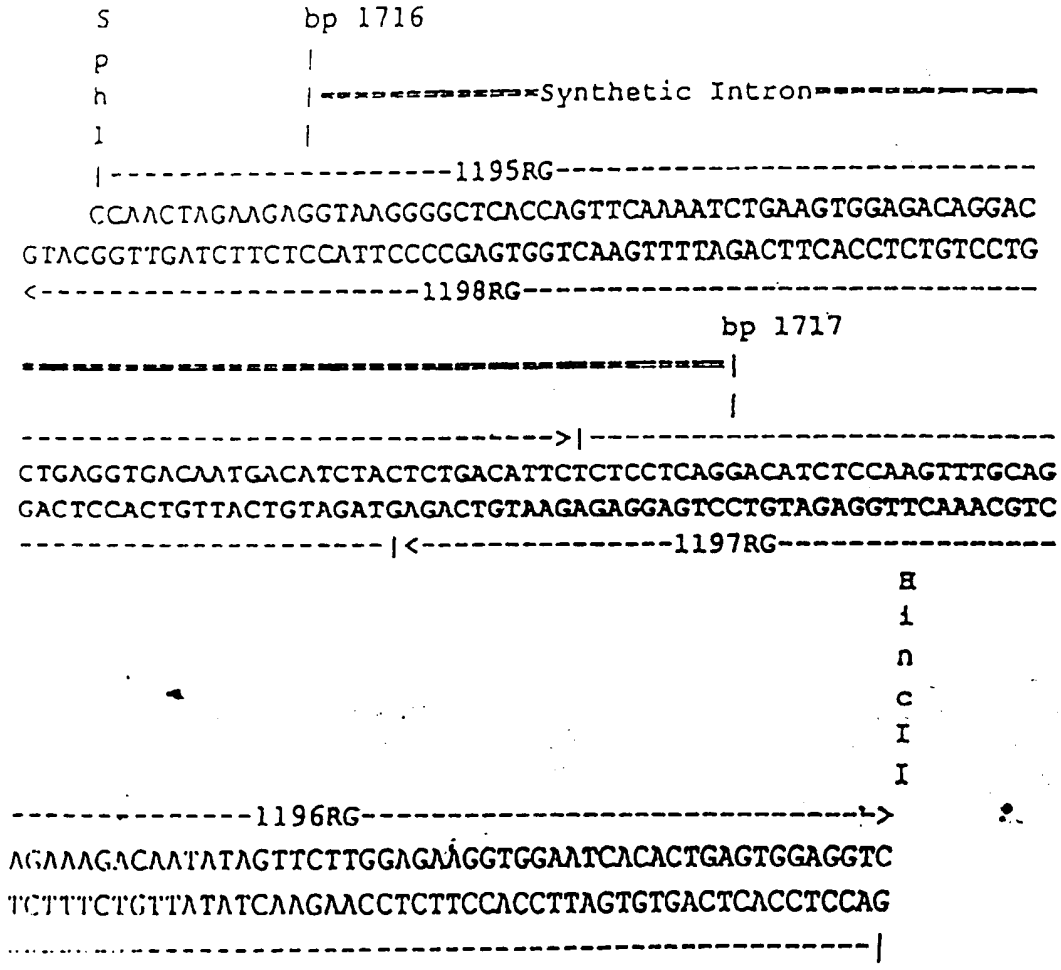

S
p
h
I
bp 1716
Synthetic Intron
1195RG
CCAACTAGAAGAGGTAAGGGGCTCACCAGTTCAAAATCTGAAGTGGAGACAGGAC
GTACGGTTGATCTTCTCCATTCCCCGAGTGGTCAAGTTTTAGACTTCACCTCTGTCCTG
1198RG
bp 1717
CTGAGGTGACAATGACATCTACTCTGACATTCTCTCCTCAGGACATCTCCAAGTTTGCAG
GACTCCACTGTTACTGTAGATGAGACTGTAAGAGAGGAGTCCTGTAGAGGTTCAAACGTC
1197RG
H
i
n
c
I
I
1196RG
AGAAAGACAATATAGTTCTTGGAGAAGGTGGAATCACACTGAGTGGAGGTC
TCTTTCTGTTATATCAAGAACCTCTTCCACCTTAGTGTGACTCACCTCCAG

Figure 6

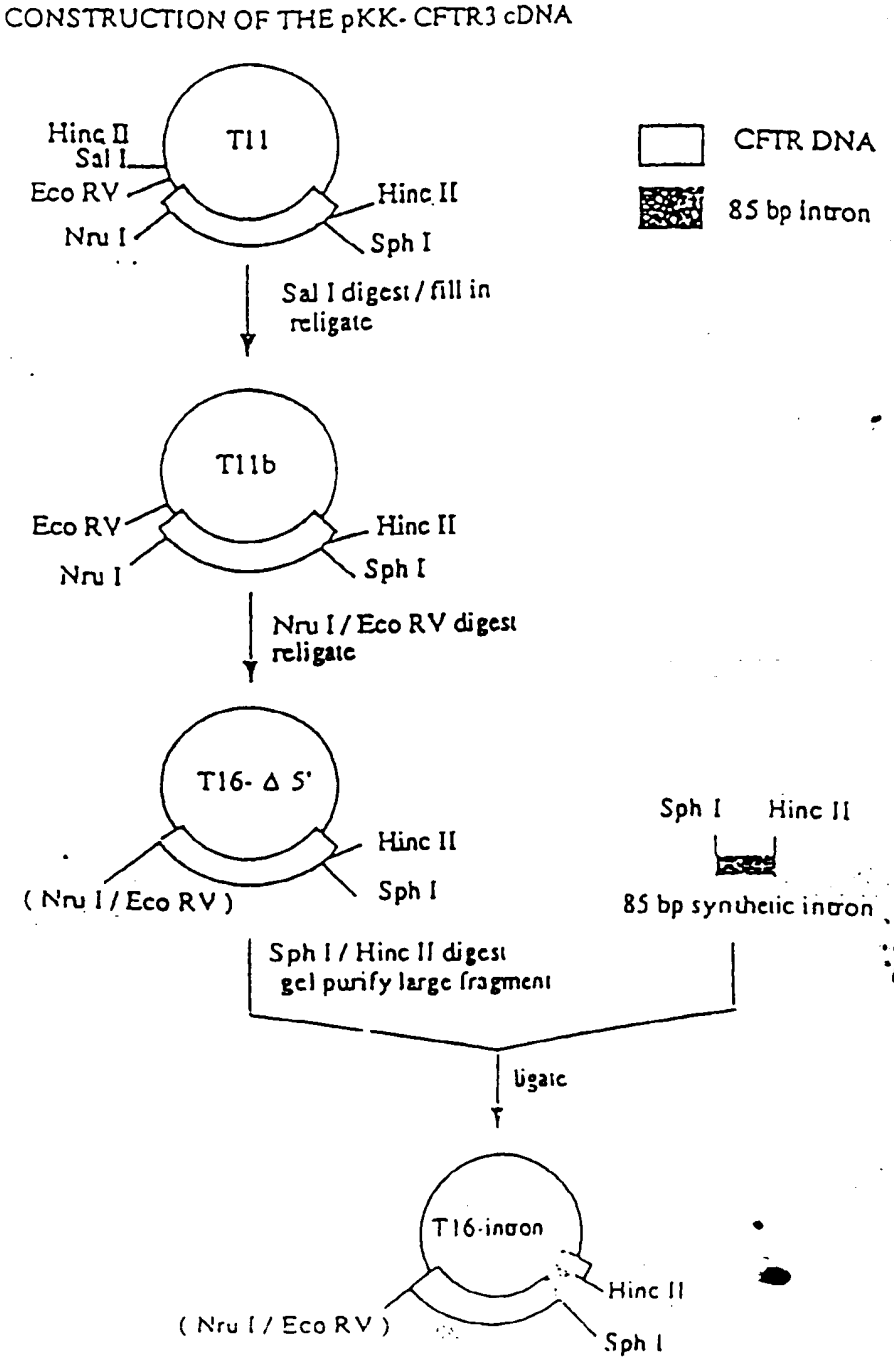

CONSTRUCTION OF THE pKK- CFTR3 cDNA
T11
Hinc II
Sal I
Eco RV
Nru I
Hinc II
Sph I
CFTR DNA
85 bp intron
Sal I digest / fill in
religate
T11b
Eco RV
Nru I
Hinc II
Sph I
Nru I / Eco RV digest
religate
T16- Δ 5'
Hinc II
(Nru I / Eco RV)
Sph I
Sph I
Hinc II
85 bp synthetic intron
Sph I / Hinc II digest
gel purify large fragment
ligate
T16-intron
Hinc II
(Nru I / Eco RV)
Sph I

Figure 7A

CONSTRUCTION OF THE pKK- CFTR3 CLONE (cont'd.)

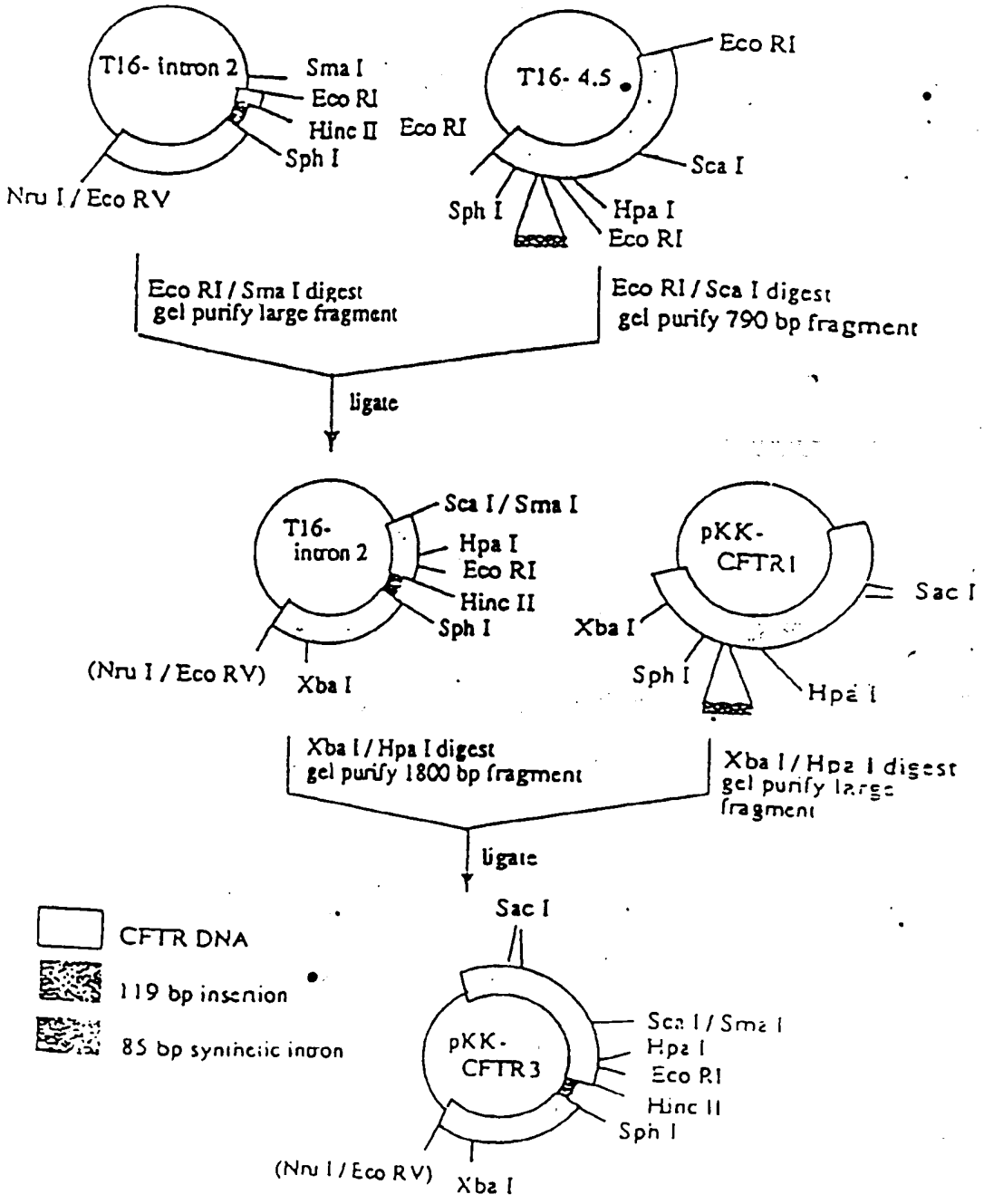


Figure 7B

MAP OF pKK- CFTR3

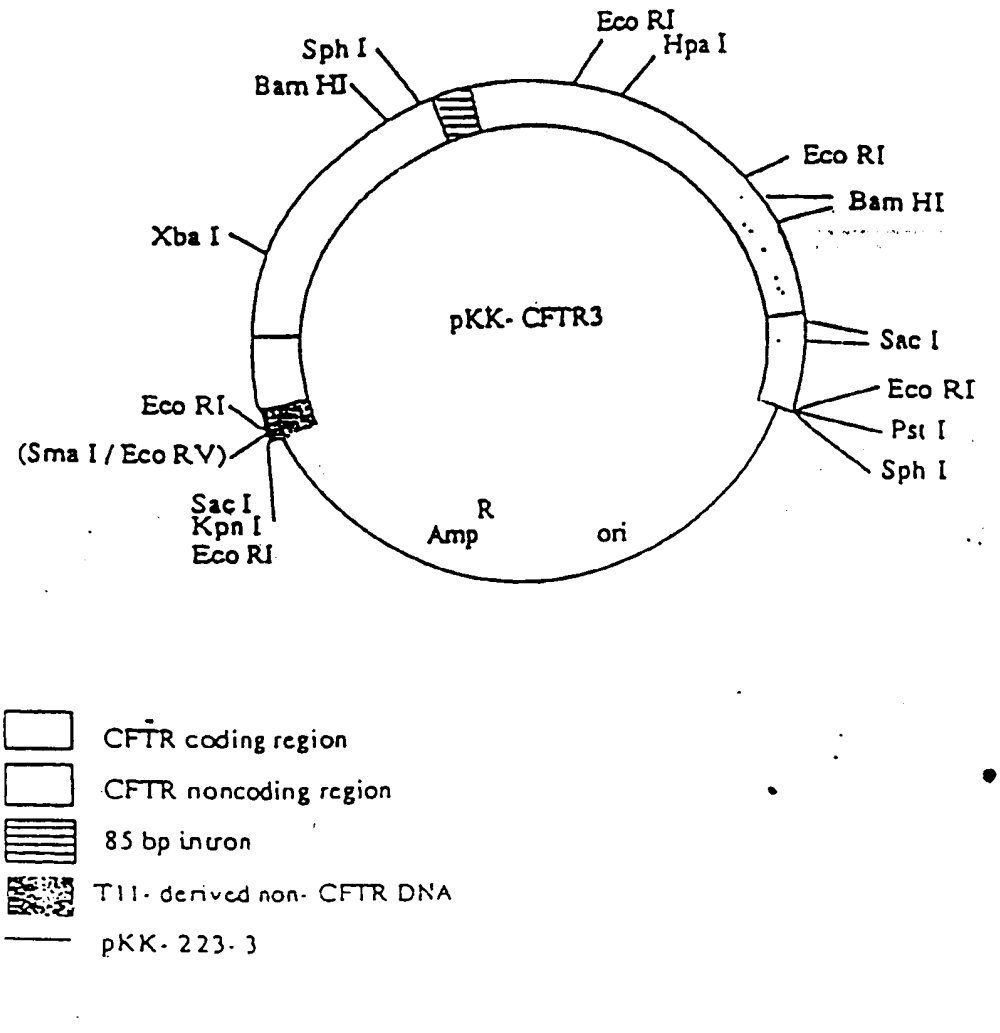


Figure 8

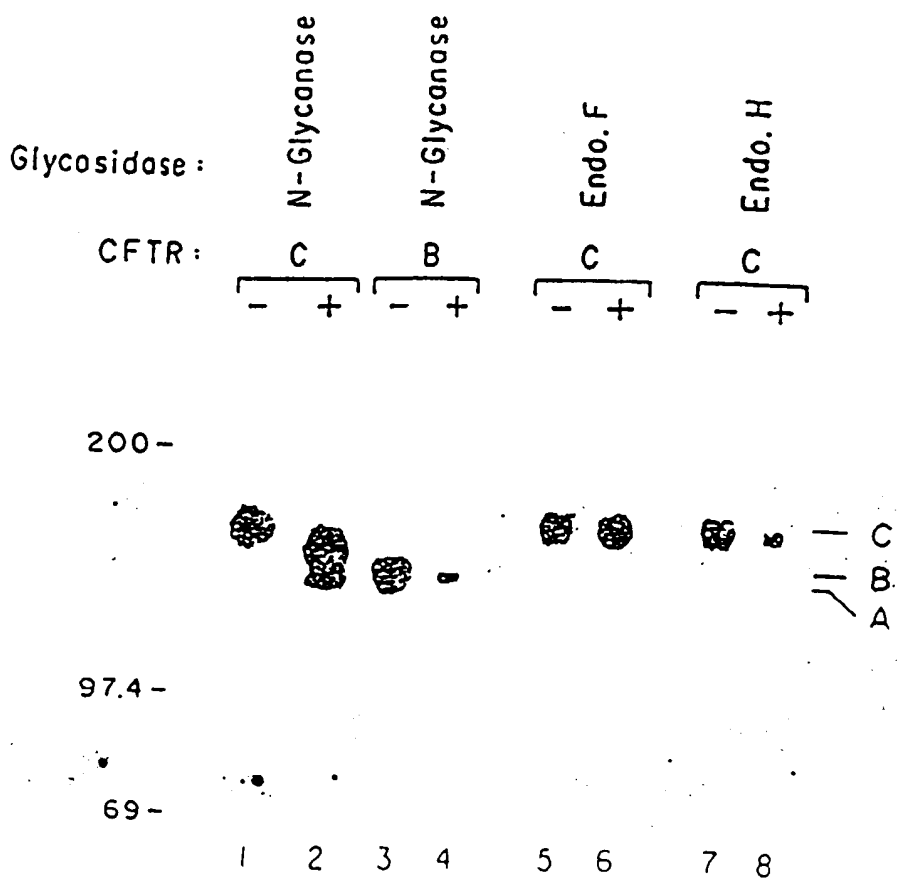

Glycosidase :
N-Glycanase
N-Glycanase
Endo.F
Endo.H
CFTR :
C
B
C
C
- +
- +
- +
- +
200-
97.4-
69-
C
B
A
1 2 3 4 5 6 7 8

Figure 9

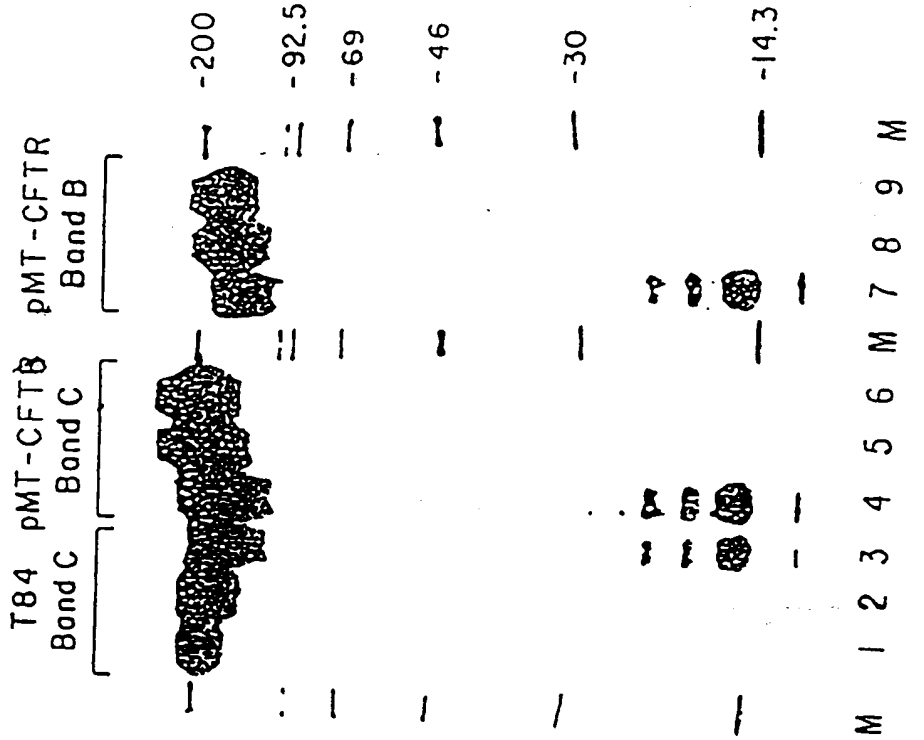


Figure 10B

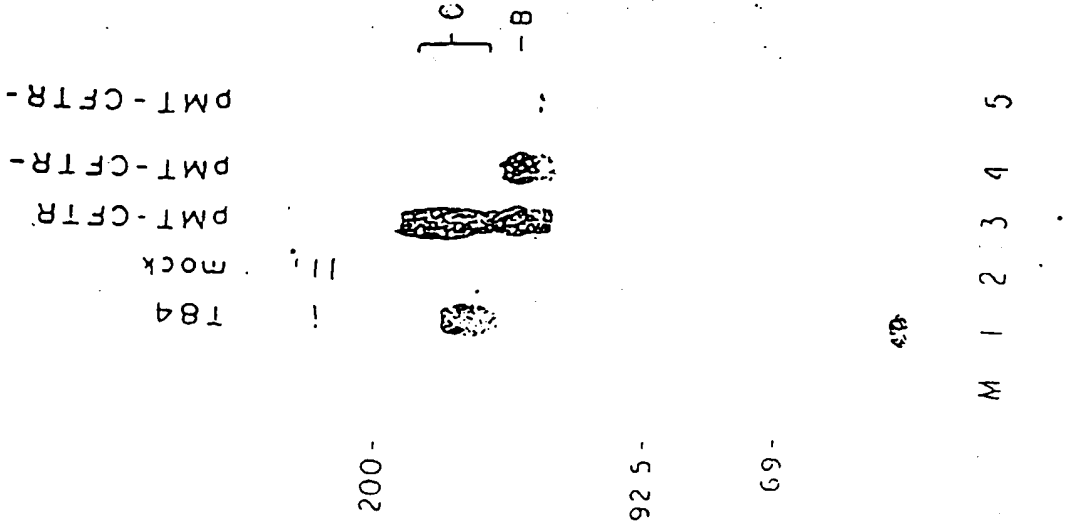


Figure 10A

Figure 11B

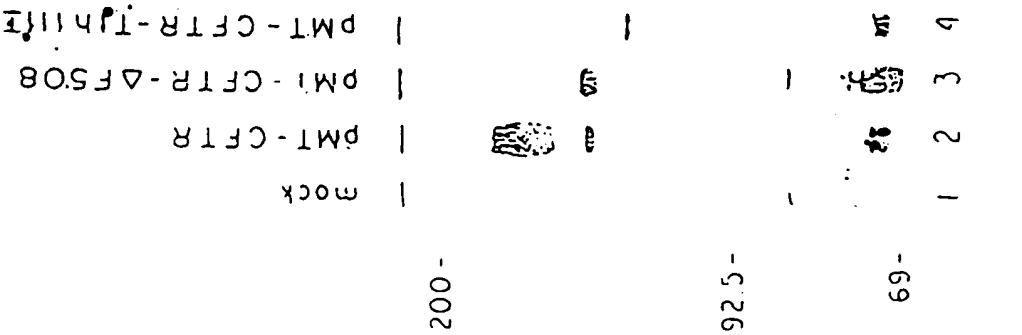


Figure 11A

Figure 12A

Figure 12B

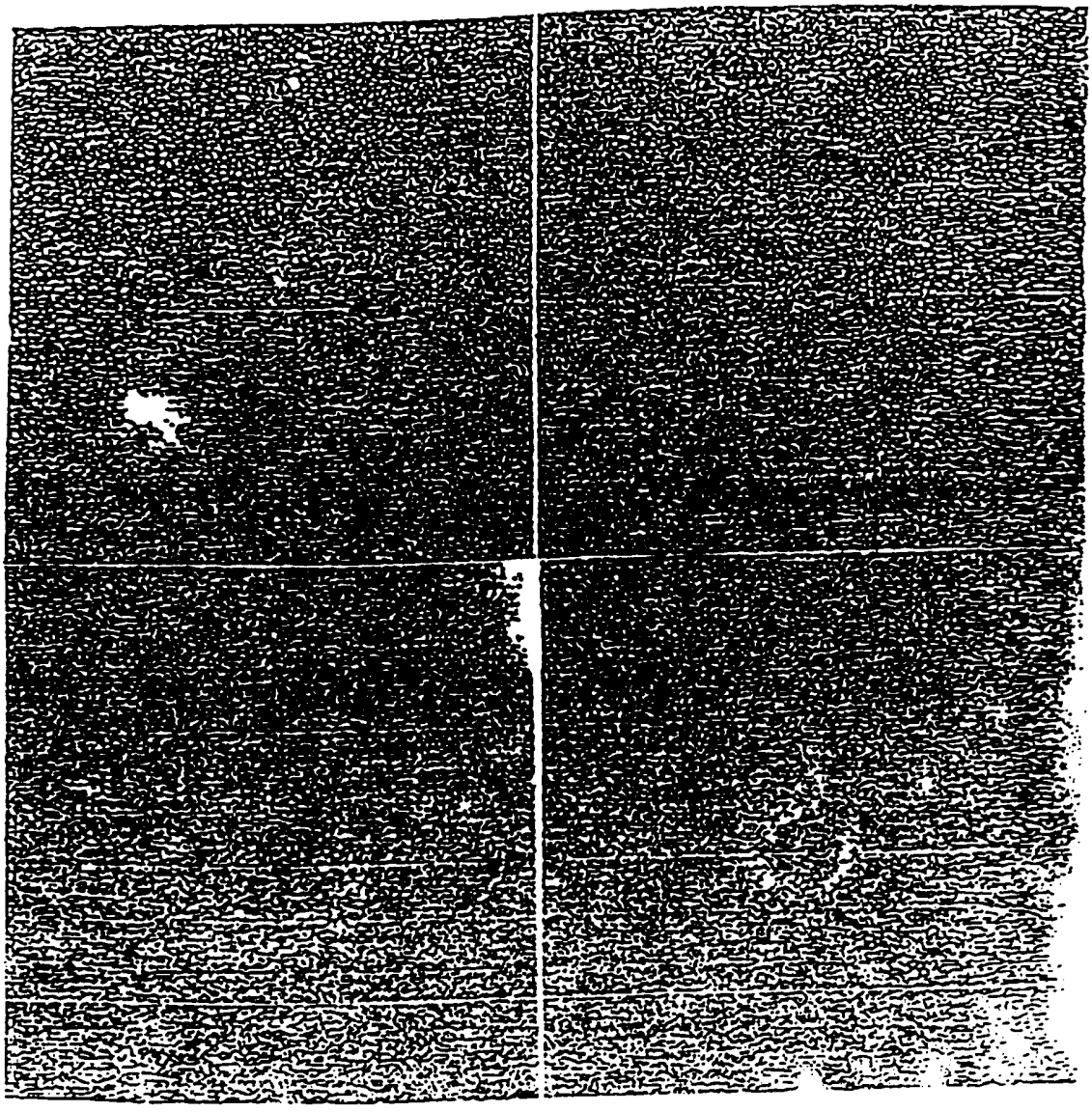


Figure 12C

Figure 12D

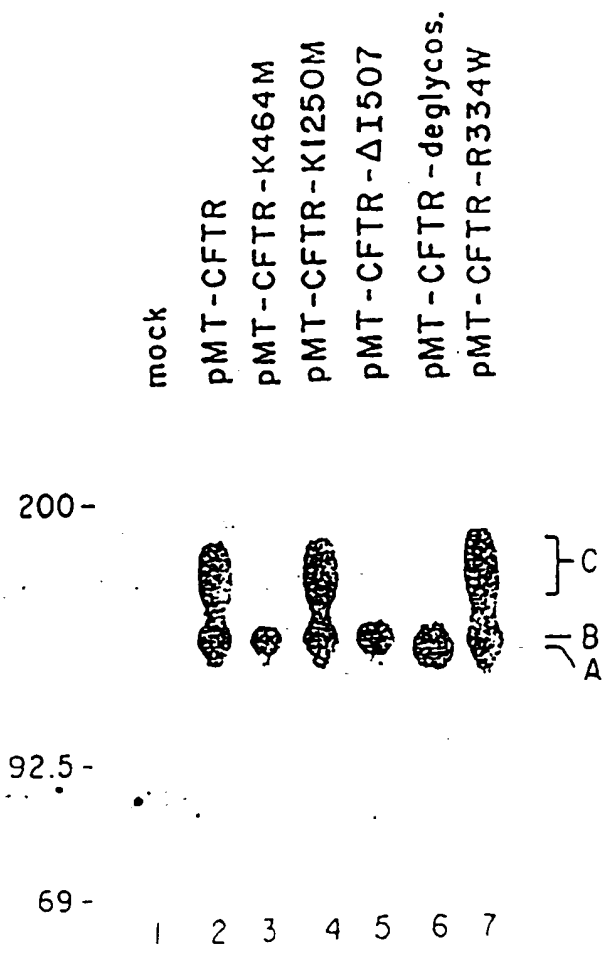

mock
pMT-CFTR
pMT-CFTR-K464M
pMT-CFTR-K1250M
pMT-CFTR-ΔI507
pMT-CFTR-deglycos.
pMT-CFTR-R334W
200-
92.5-
69-
1 2 3 4 5 6 7
C
B
A

Figure 13

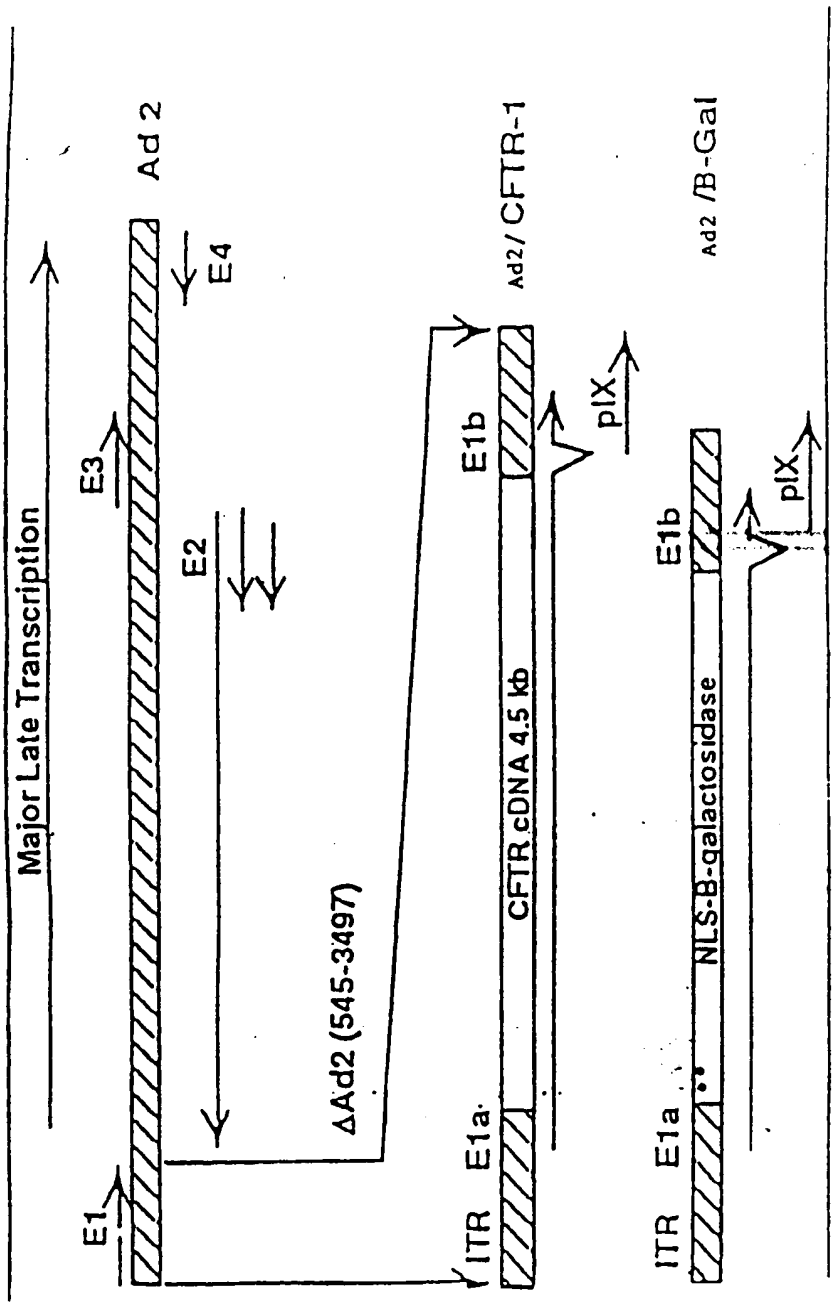


FIGURE 1
MAP OF VECTOR
Major Late Transcription
Ad 2
E4
E3
E2
E1
ΔAd2 (545-3497)
Ad2/ CFTR-1
E1b
pIX
CFTR cDNA 4.5 kb
ITR
E1a
Ad2 /B-Gal
E1b
pIX
NLS-B-galactosidase
ITR
E1a

Figure 14

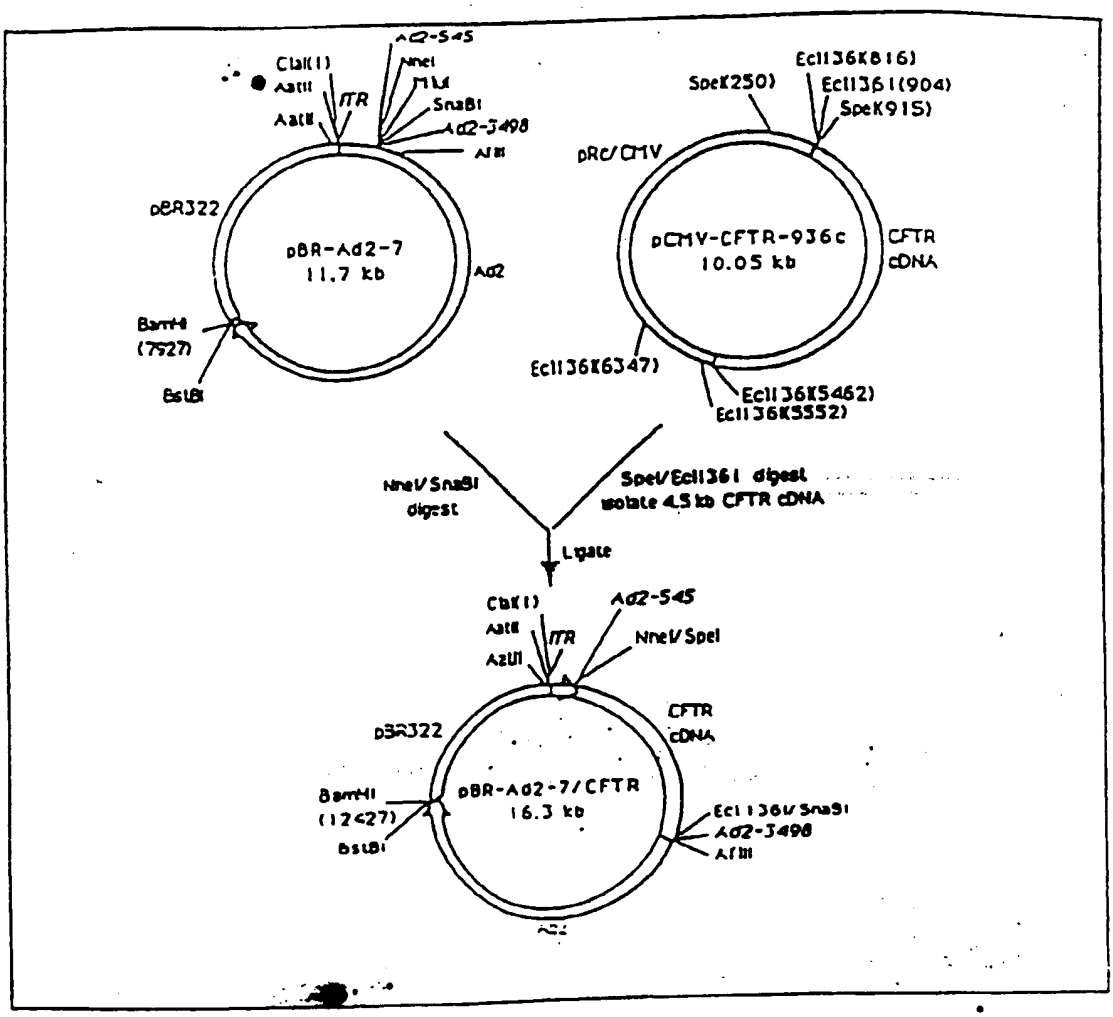

ClaI(1)
AatII
ITR
Ad2-545
NheI
MluI
SnaBI
Ad2-3498
AflII
AatII
pBR322
pBR-Ad2-7
11.7 kb
Ad2
BamHI
(7527)
BstBI
pRc/CMV
SpeI(250)
Ec1136I(816)
Ec1136I(904)
SpeI(915)
pCMV-CFTR-936c
10.05 kb
CFTR
cDNA
Ec1136I(6347)
Ec1136I(5462)
Ec1136I(5552)
NheI/SnaBI
digest
SpeI/Ec1136I digest
isolate 4.5 kb CFTR cDNA
Ligate
ClaI(1)
AatII
AzlII
ITR
Ad2-545
NheI/SpeI
pBR322
CFTR
cDNA
BamHI
(1247)
BstBI
pBR-Ad2-7/CFTR
16.3 kb
Ec1136I/SnaBI
Ad2-3498
AflII

Figure 15

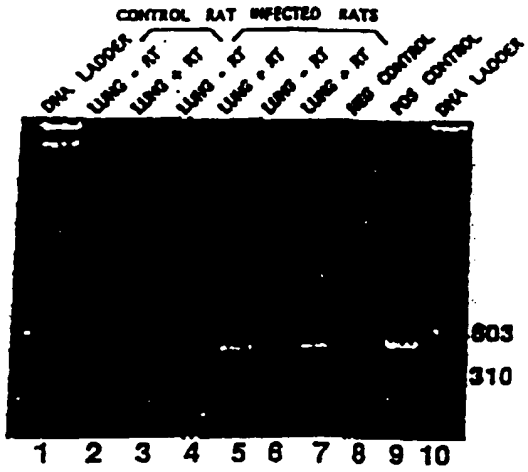

CONTROL RAT
INFECTED RATS
DNA LADDER
LUNG + RT
LUNG + RT
LUNG + RT
LUNG + RT
LUNG + RT
LUNG + RT
NEG CONTROL
POS CONTROL
DNA LADDER
603
310
1 2 3 4 5 6 7 8 9 10

Figure 16

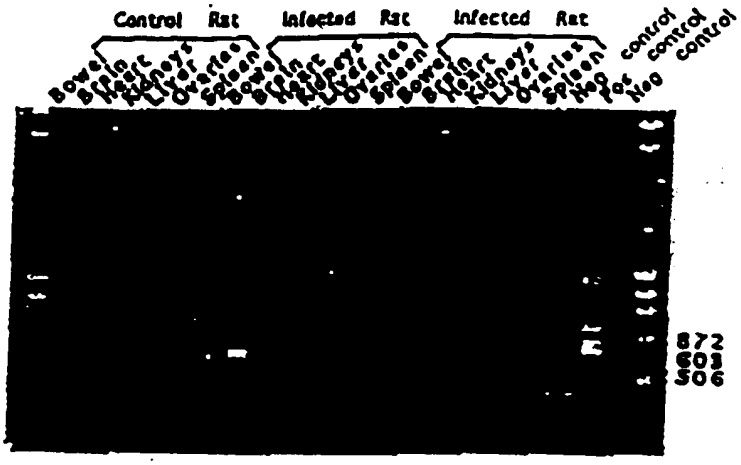

Control Rat
Infected Rat
Infected Rat
Bowel
Brain
Heart
Kidneys
Liver
Ovaries
Spleen
Bowel
Brain
Heart
Kidneys
Liver
Ovaries
Spleen
Bowel
Brain
Heart
Kidneys
Liver
Ovaries
Spleen
control
control
control
Neg
Pos
Neg
872
603
506

Figure 17

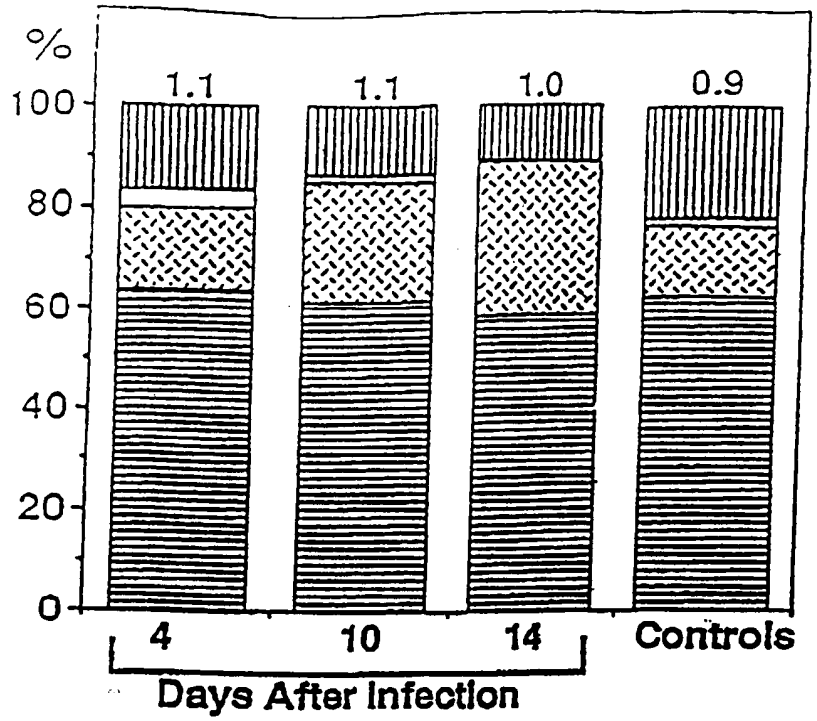


Figure 18A

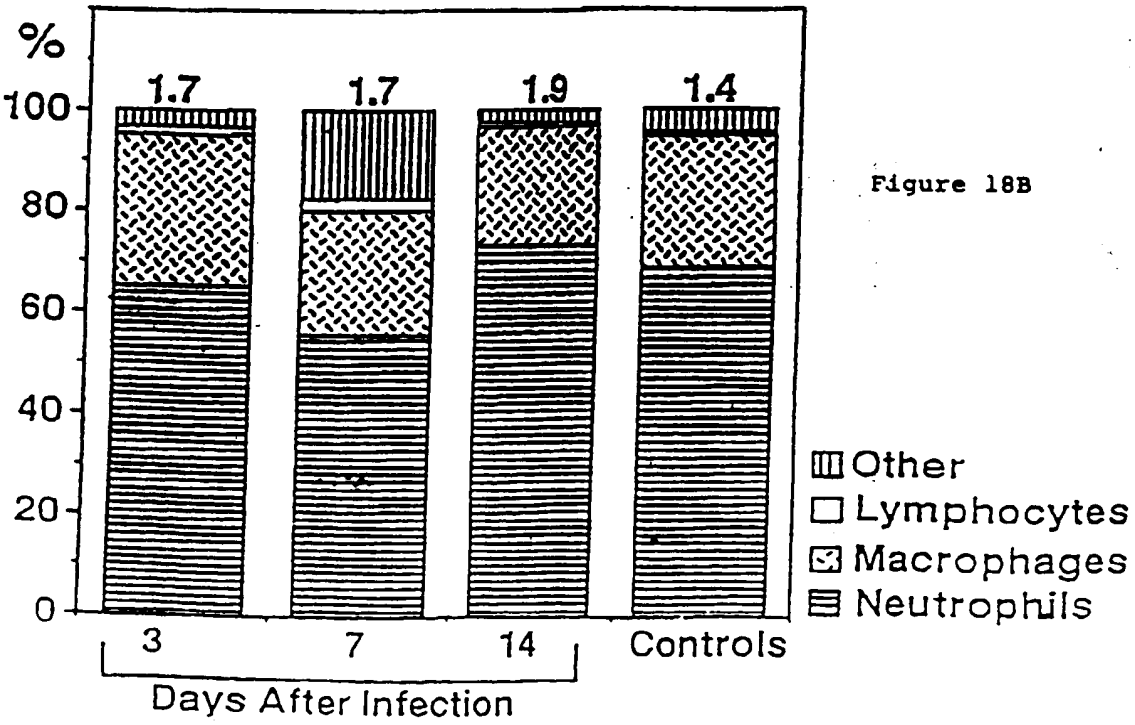


Figure 18B

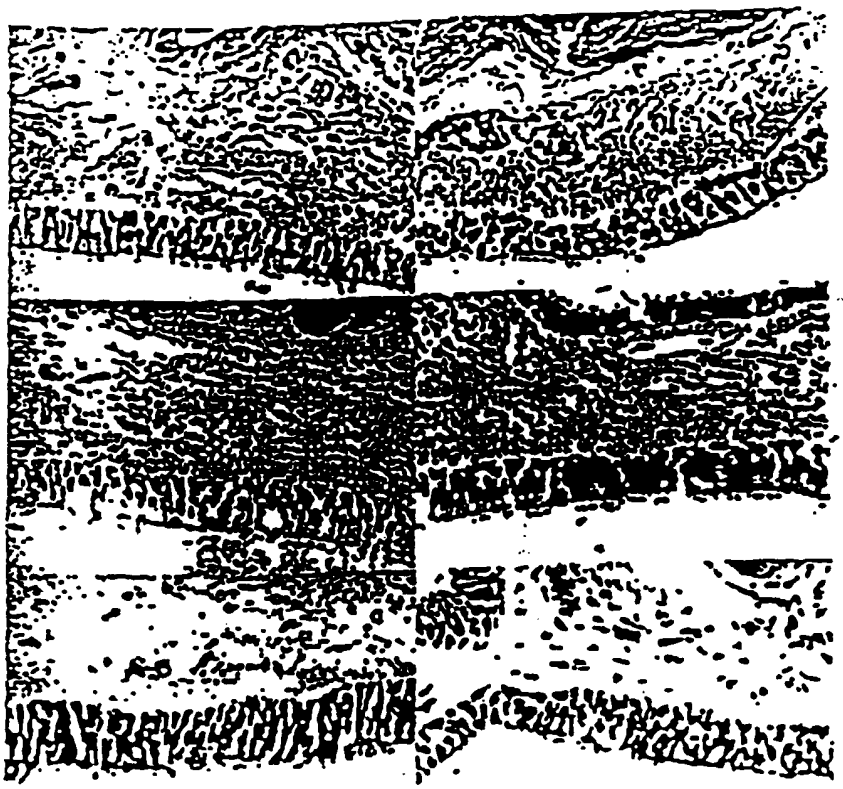


Figure 19

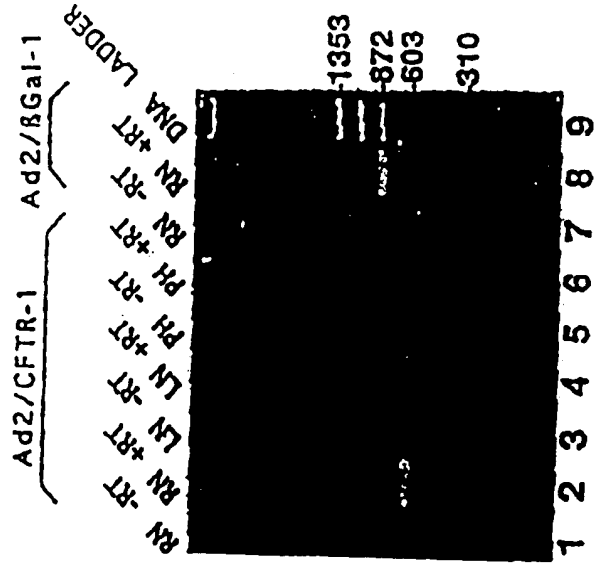


Figure 20A

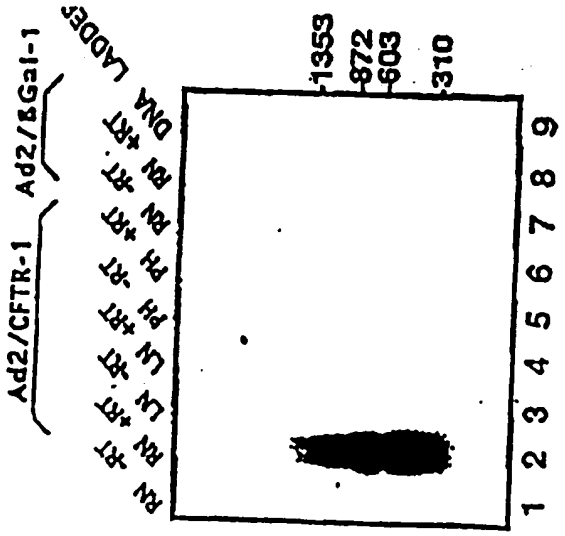


Figure 20B

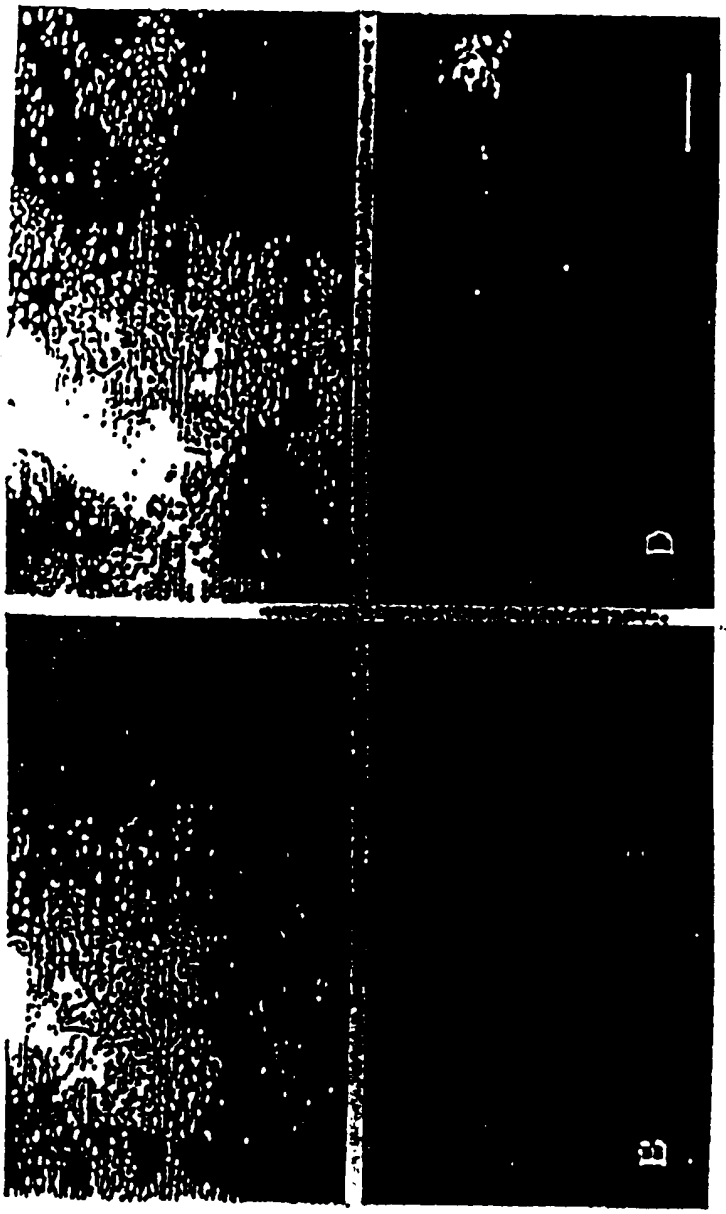


Figure 21

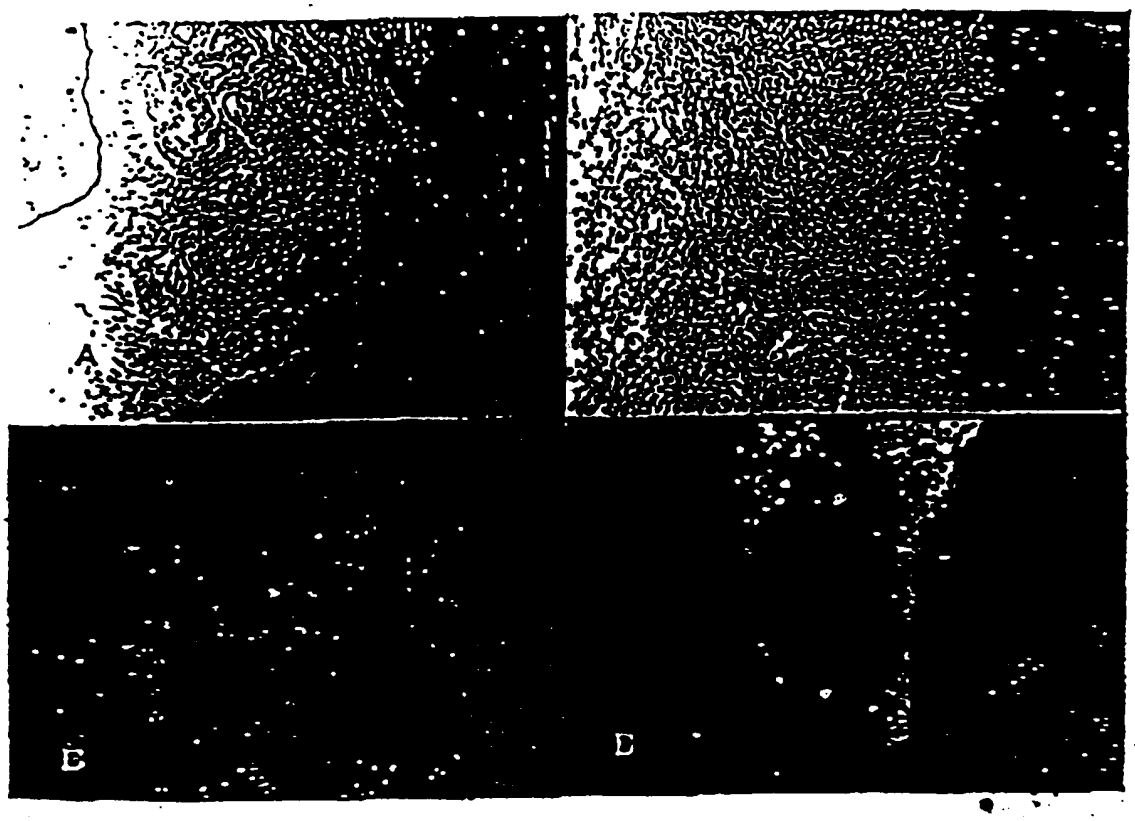


Figure 22

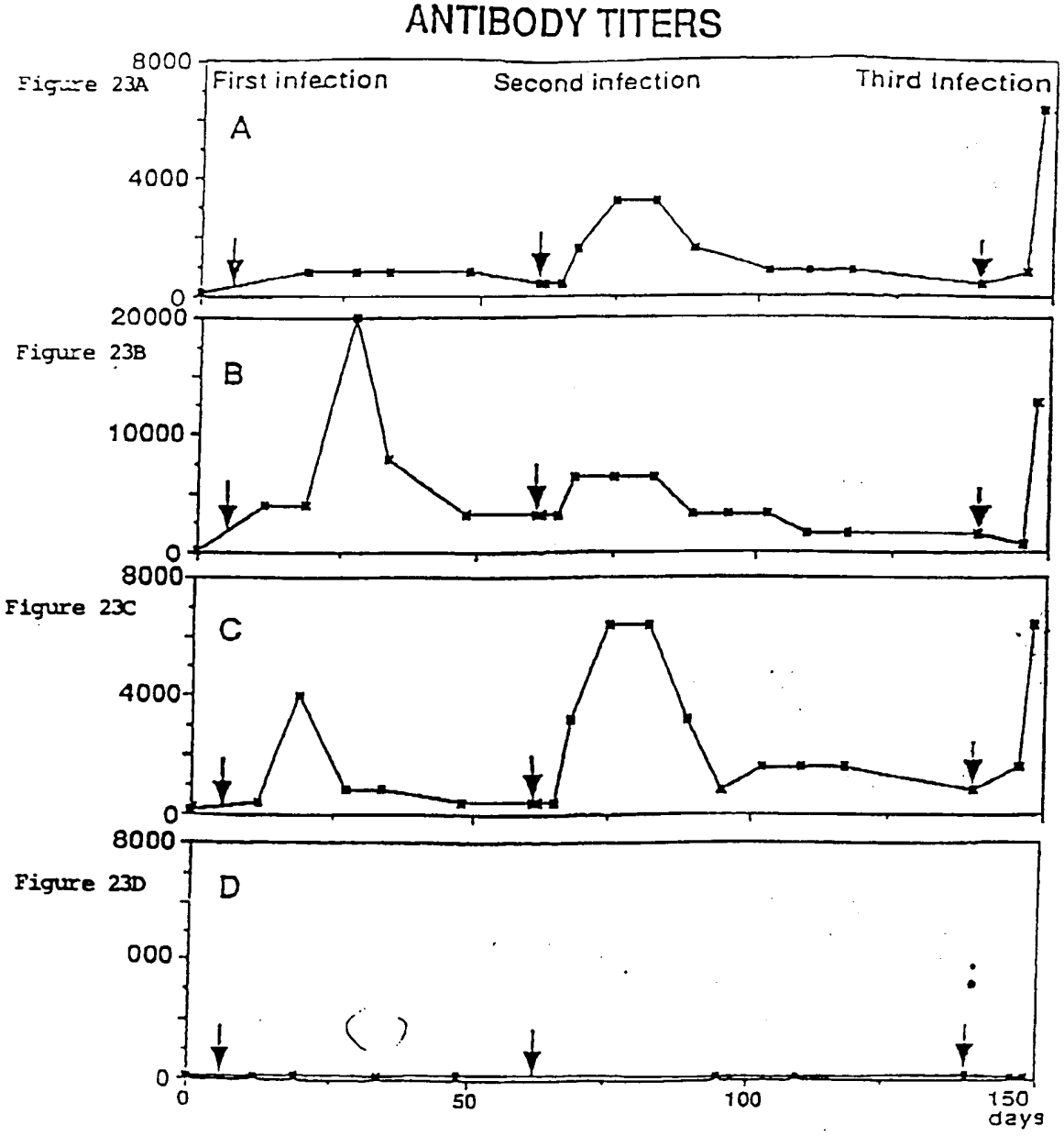


Figure 23A

Figure 23B

Figure 23C

Figure 23D

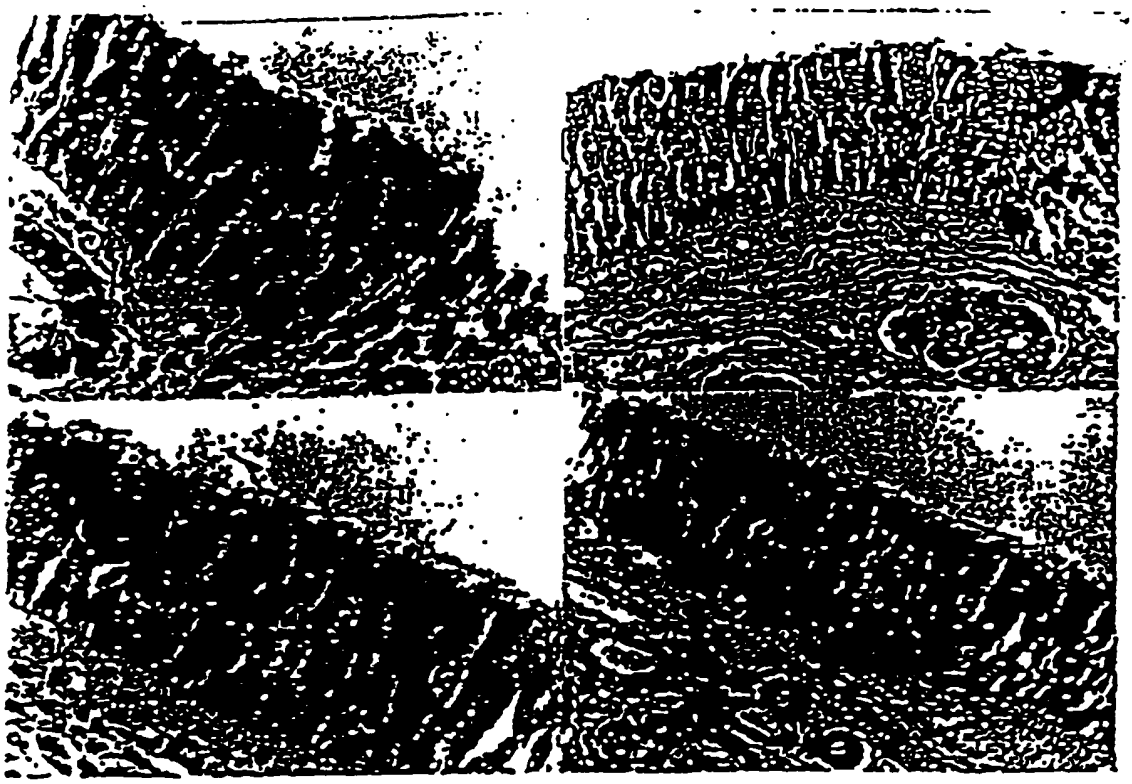


Figure 24

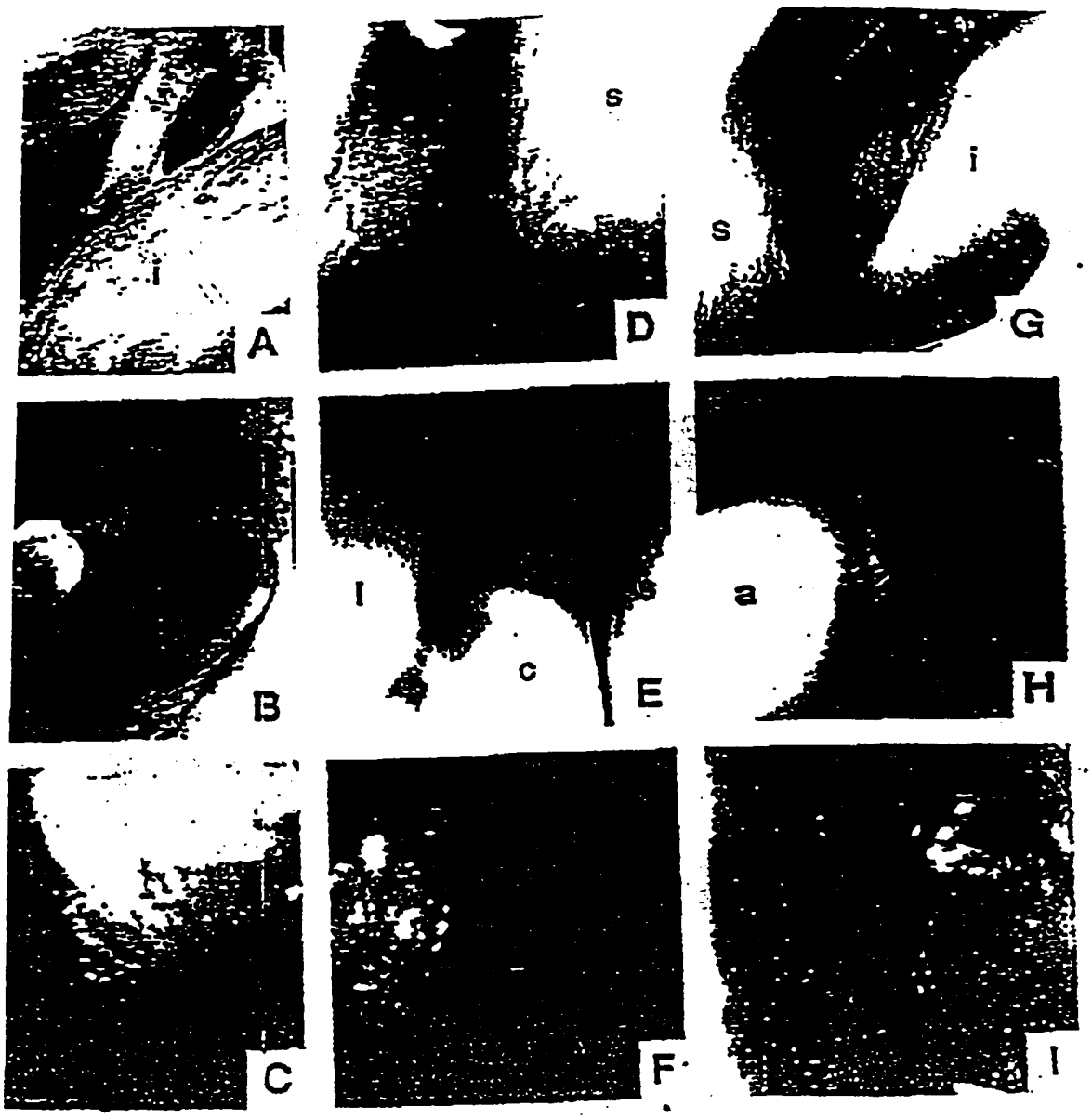


Figure 25

Figure 26

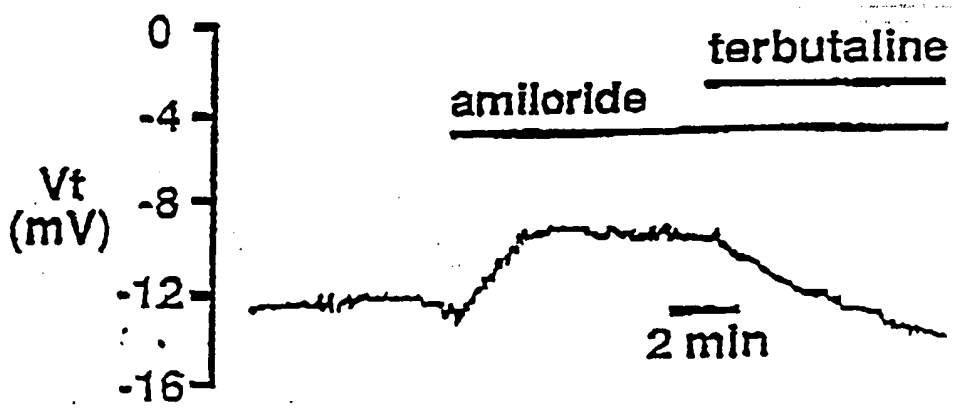


Figure 27

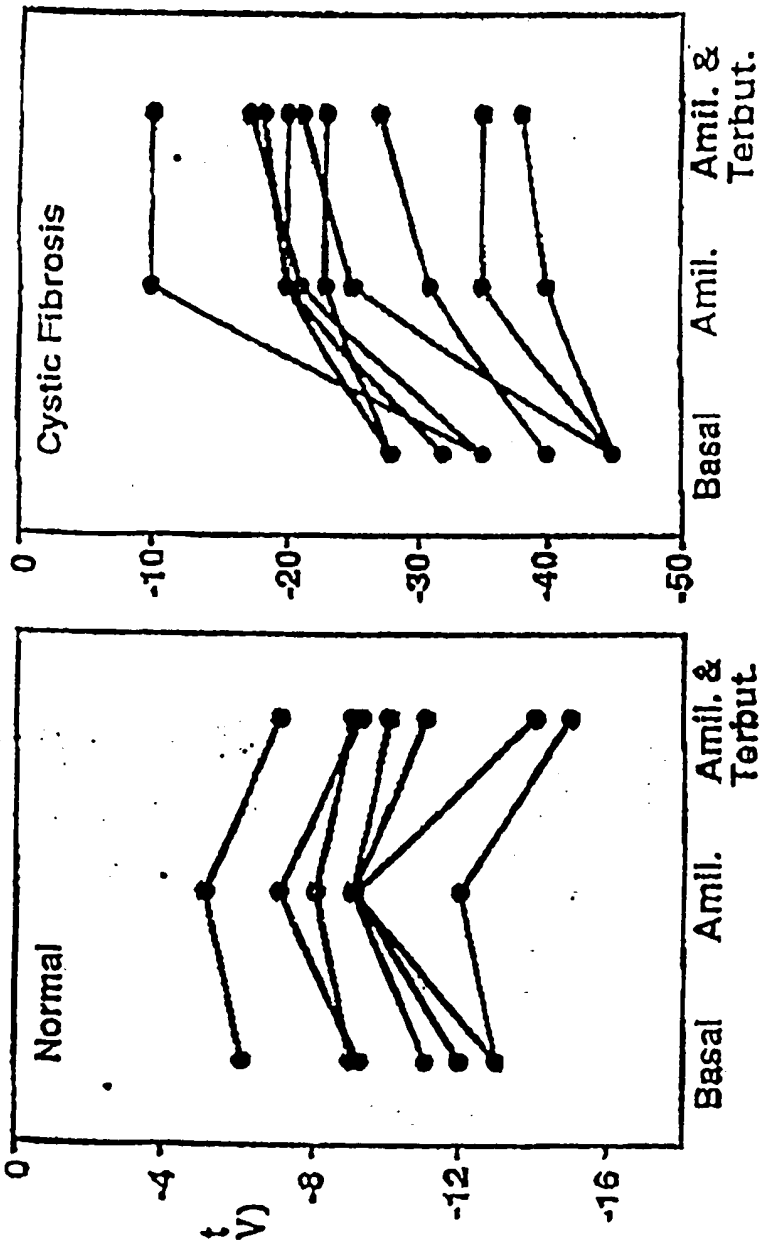


Figure 28B

Figure 28A

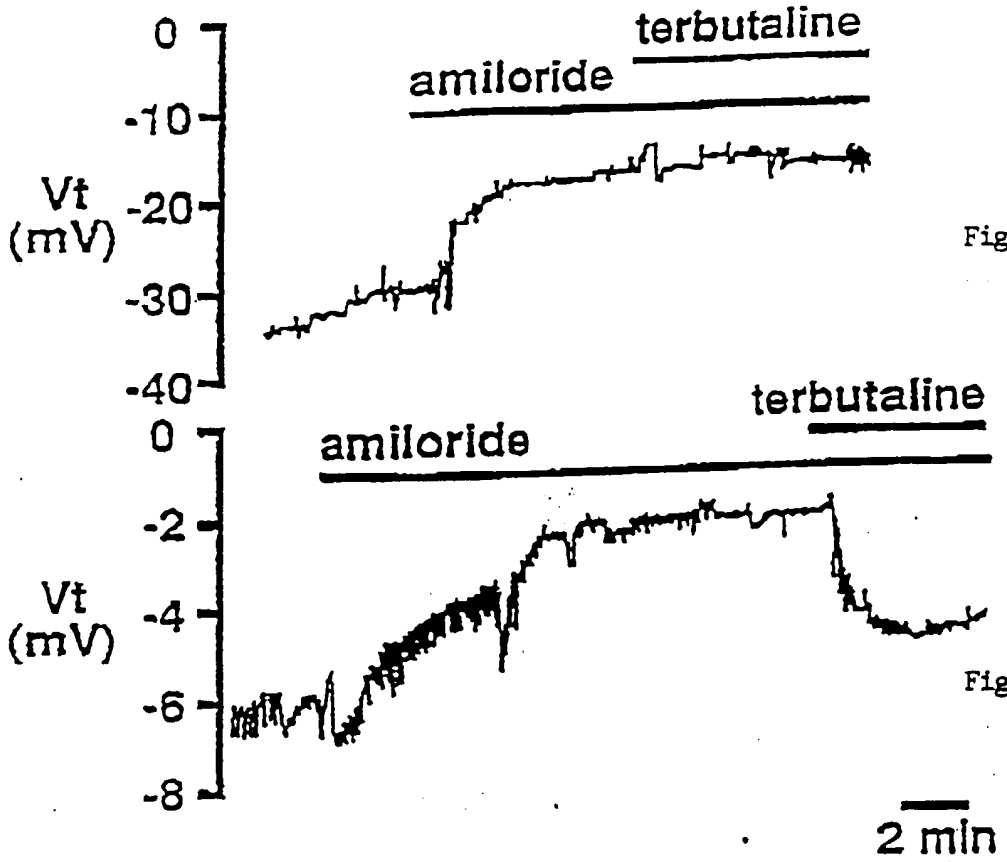


Figure 29A

Figure 29B

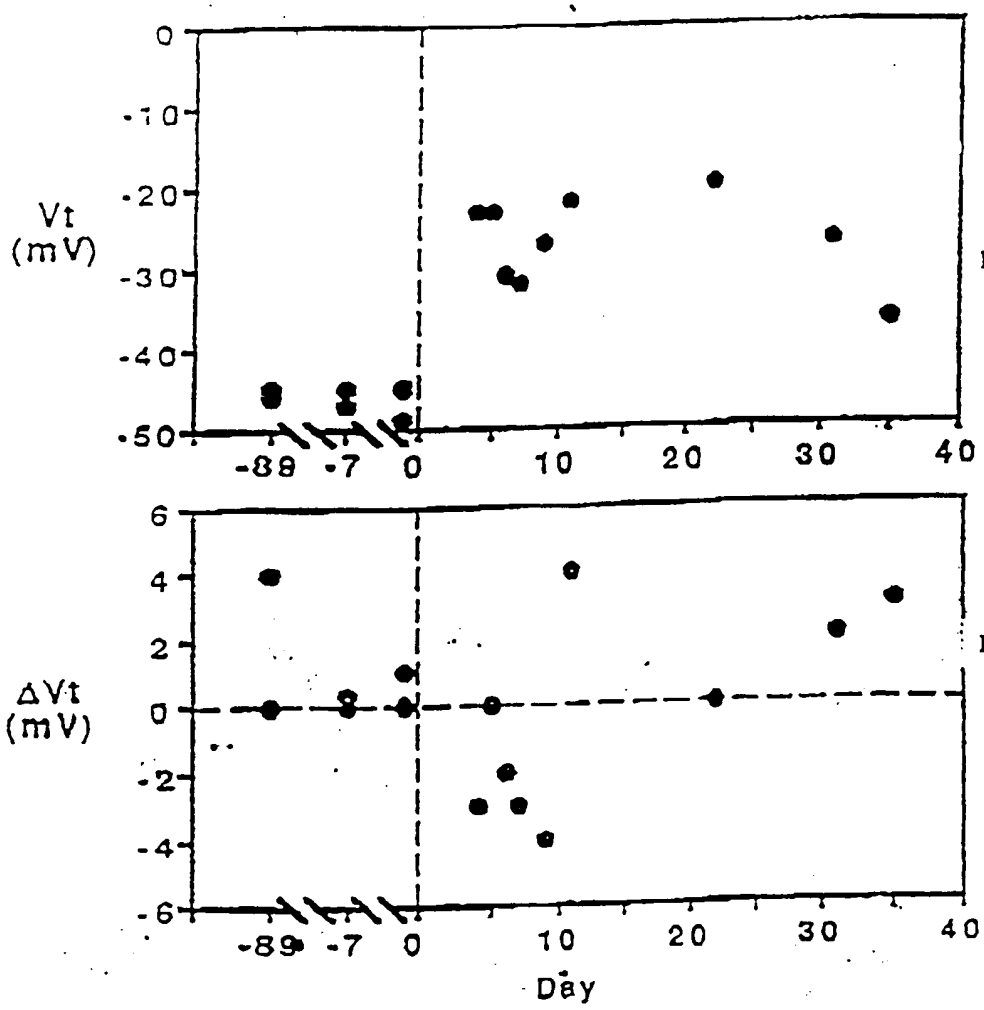


Figure 30A

Figure 30B

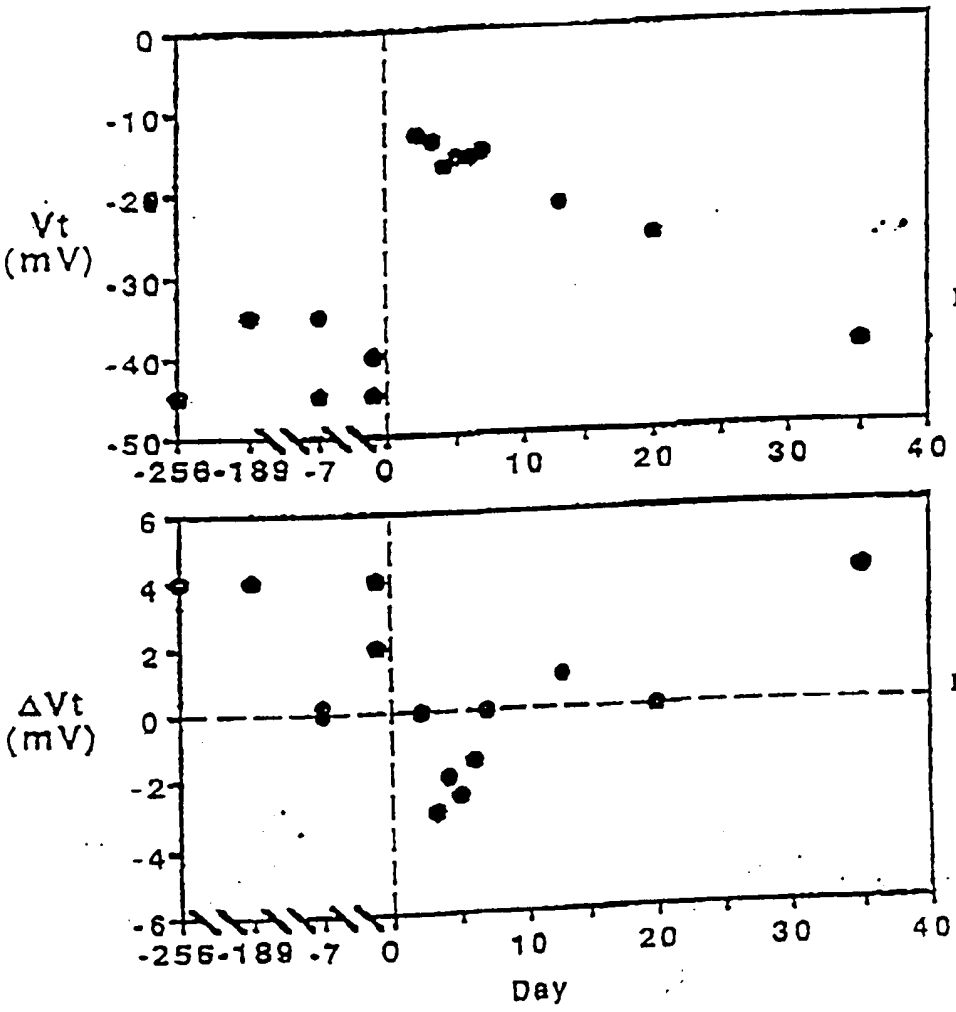


Figure 30C

Figure 30D

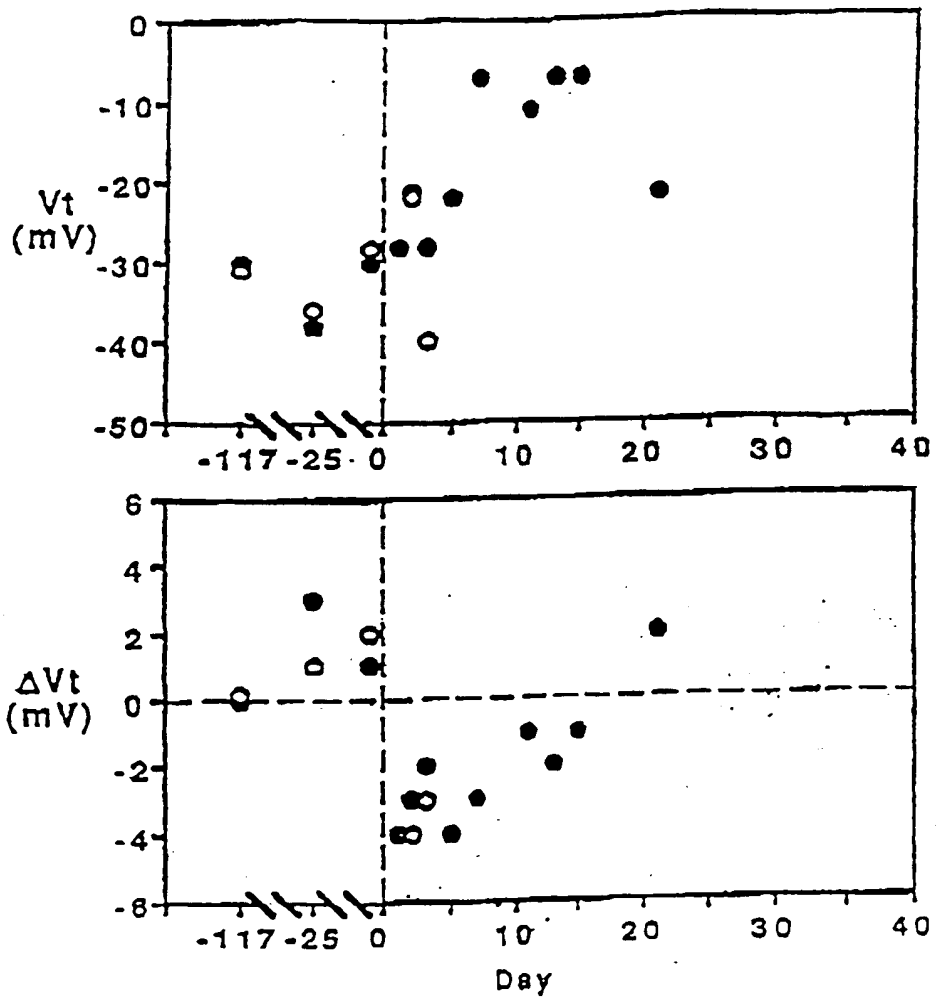


Figure 30E

Figure 30F

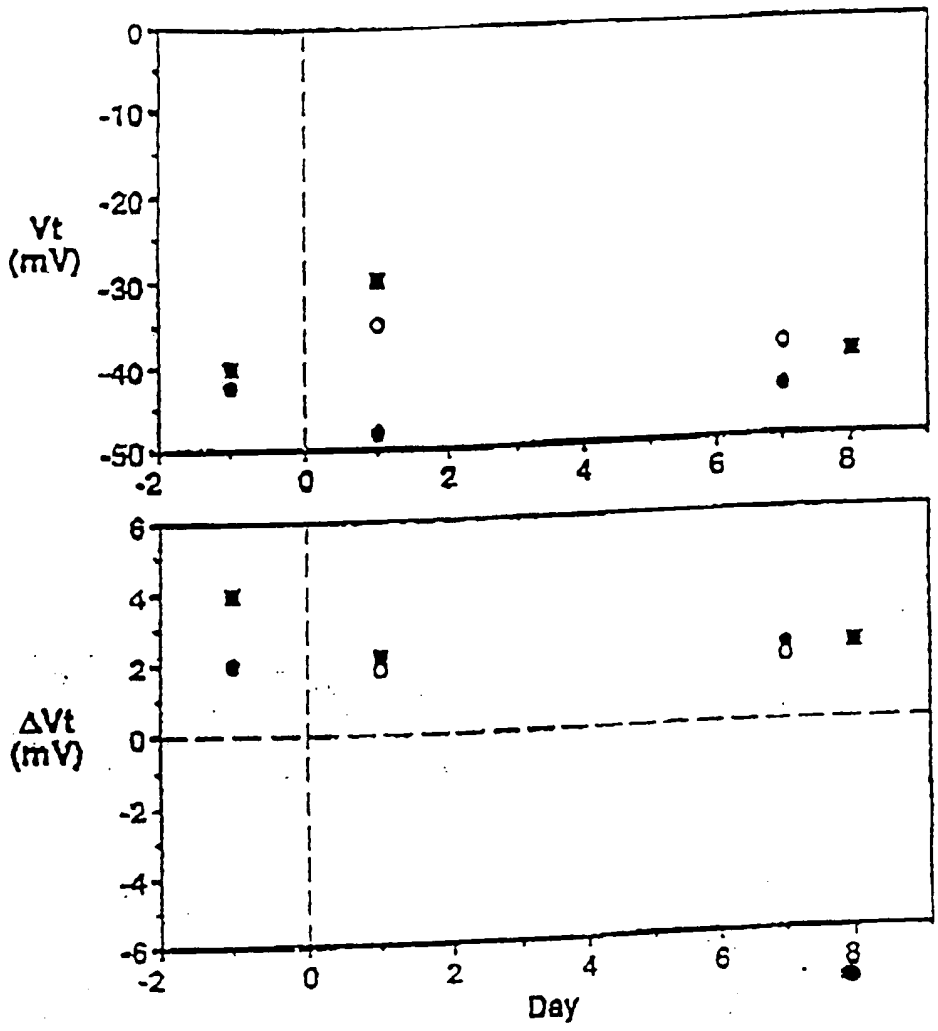
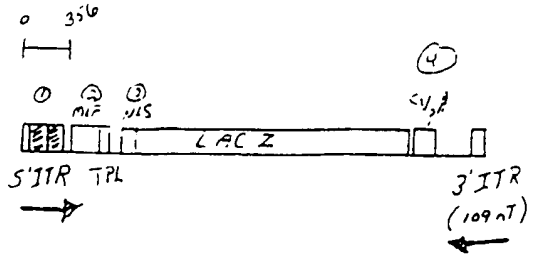
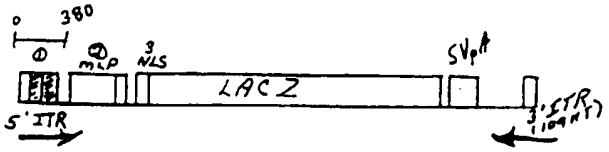


Figure 31



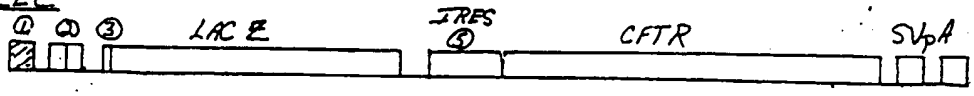
① Adenovirus Type 2 packaging signal and E1 enhancer Region
② Adenovirus Type 2 major Late Promoter and Tri-partite Leader
③ SV_{40} T-antigen nuclear Localization Signal
④ SV_{40} Poly Adenylation Signal

PAV II



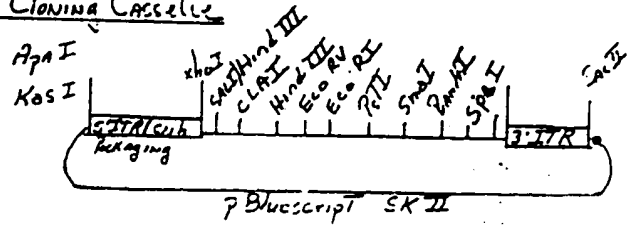
① Adenovirus Type 2 packaging signal and E1 enhancer Region
② Adenovirus Type 2 major Late Promoter and Tri-partite Leader
③ SV40 T-antigen nuclear Localization Signal
④ SV40 Poly Adenylation Signal

PAV I/II LEC



⑤ EMC Virus Internal Ribosomal entry site - for Polycistronic Translation

PAV I Cloning Cassette



Expression Cassette

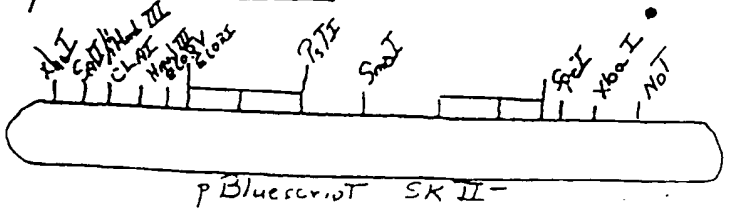


Figure 32

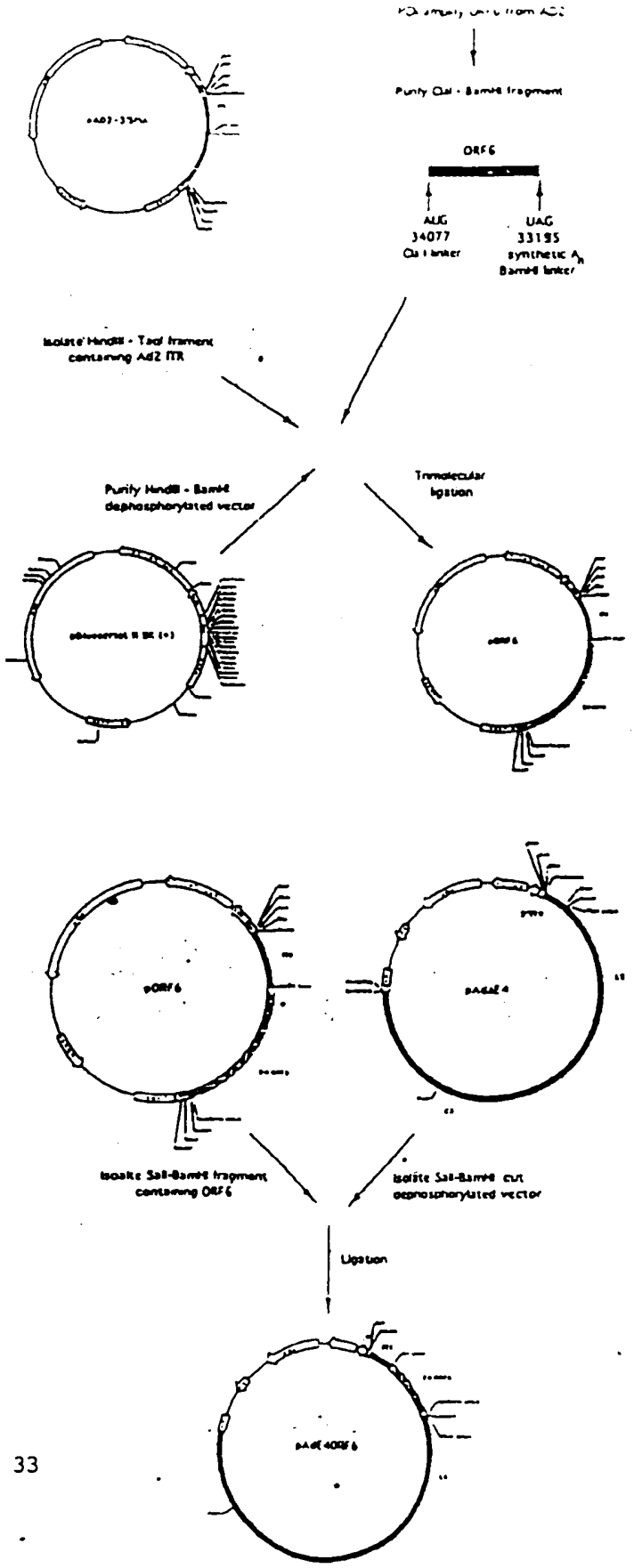

Purify Cla I - BamHI fragment
ORF6
ATG
34077
UAG
33195
synthetic A
BamHI linker
Isolate HindIII - TaqI fragment containing Ad2 ITR
Purify HindIII - BamHI dephosphorylated vector
Trimolecular ligation
pORF6
Isolate SalI-BamHI fragment containing ORF6
Isolate SalI-BamHI cut dephosphorylated vector
Ligation

Figure 33

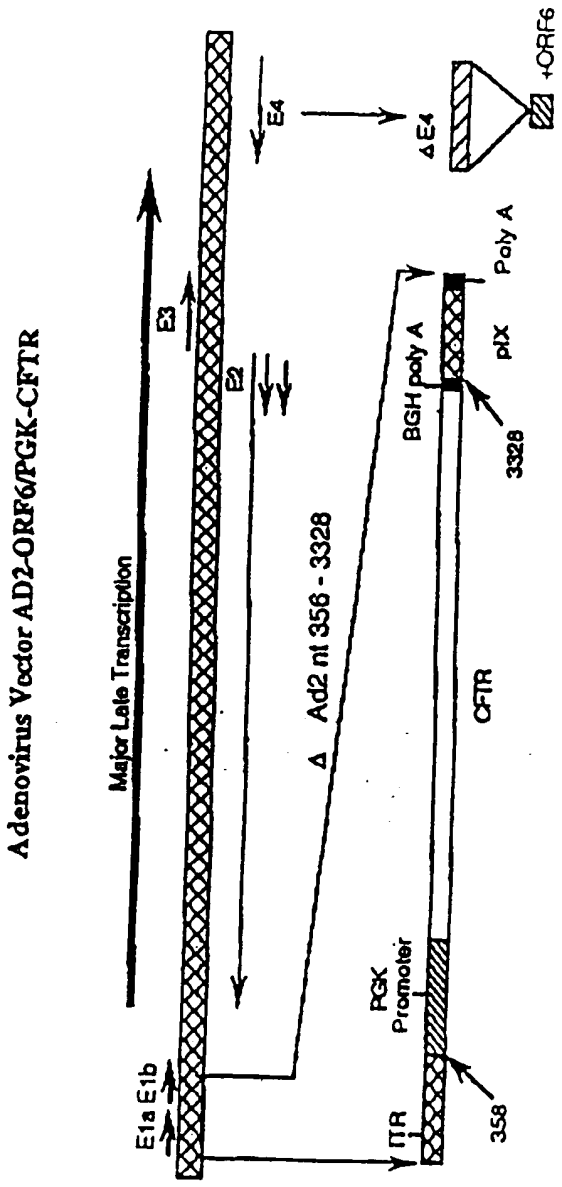

Adenovirus Vector AD2-ORF6/PGK-CFTR
Major Late Transcription
E1a E1b
E3
E2
E4
ΔE4
+ORF6
Δ Ad2 nt 356 - 3328
BGH poly A
Poly A
pIX
3328
CFTR
PGK Promoter
ITR
358

Figure 34

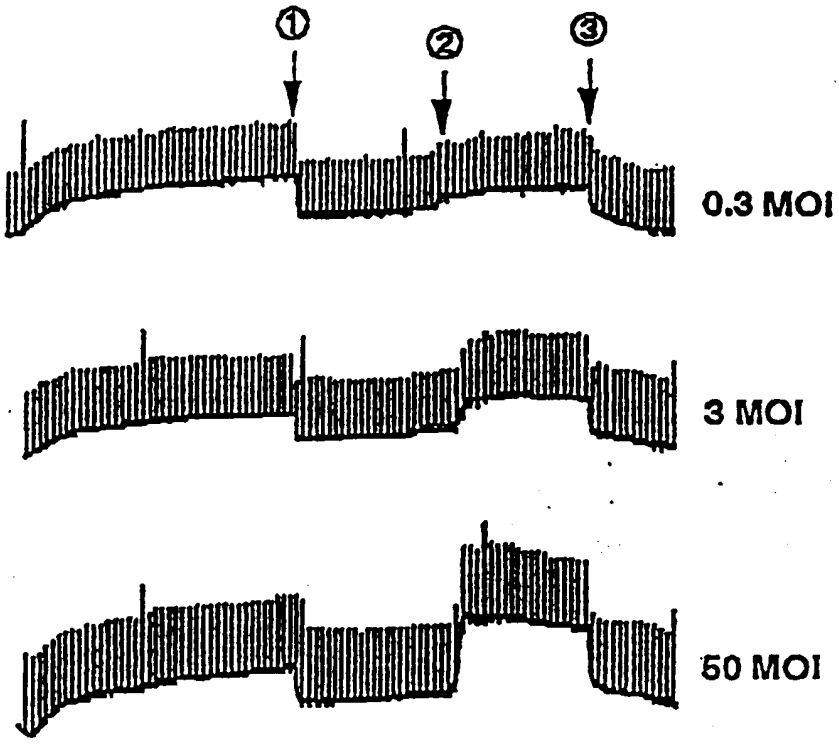


Figure 35

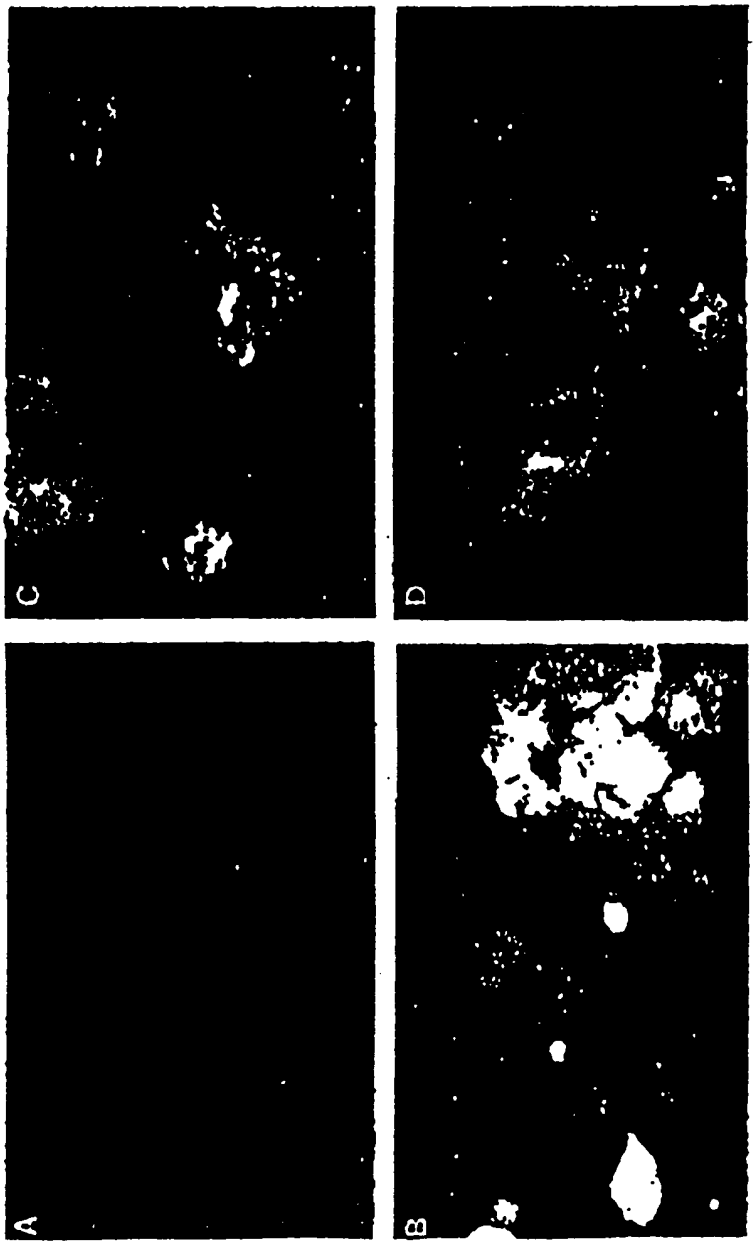


Figure 36 C

Figure 36D

Figure 36A

Figure 36B

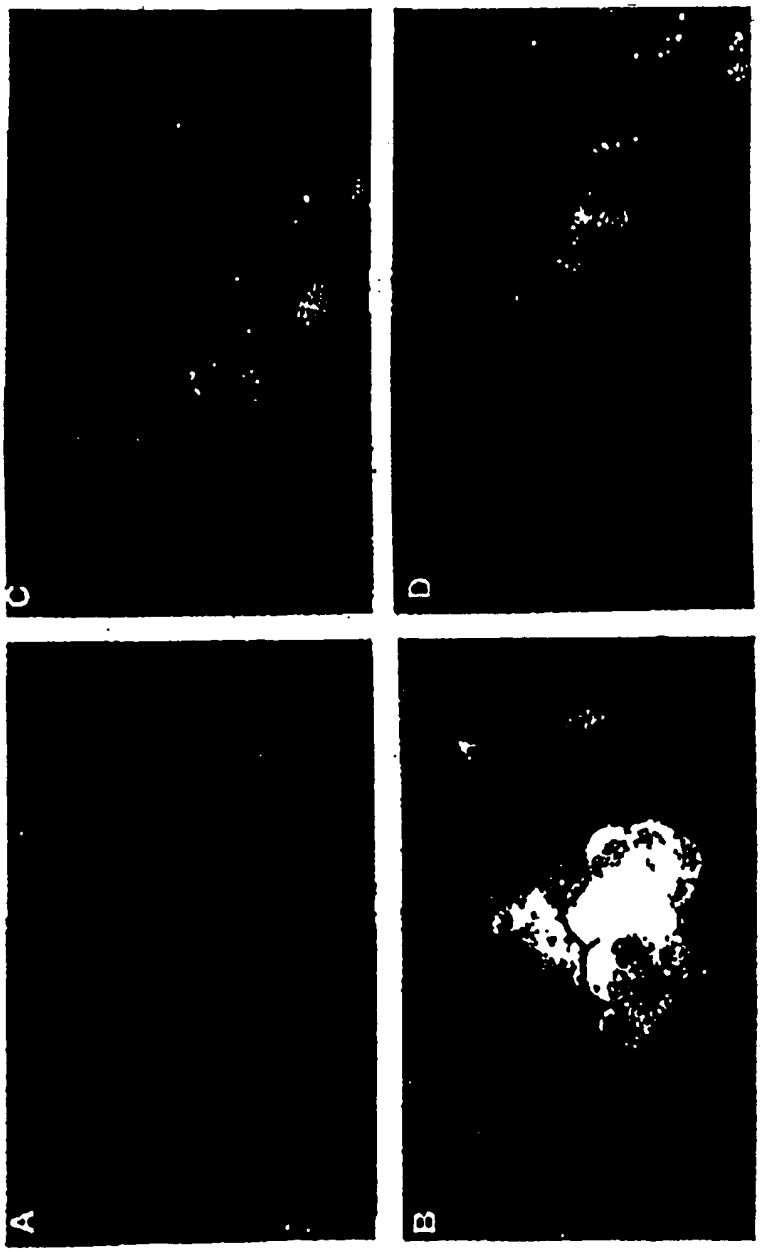


Figure 37C

Figure 37D

Figure 37A

Figure 37B

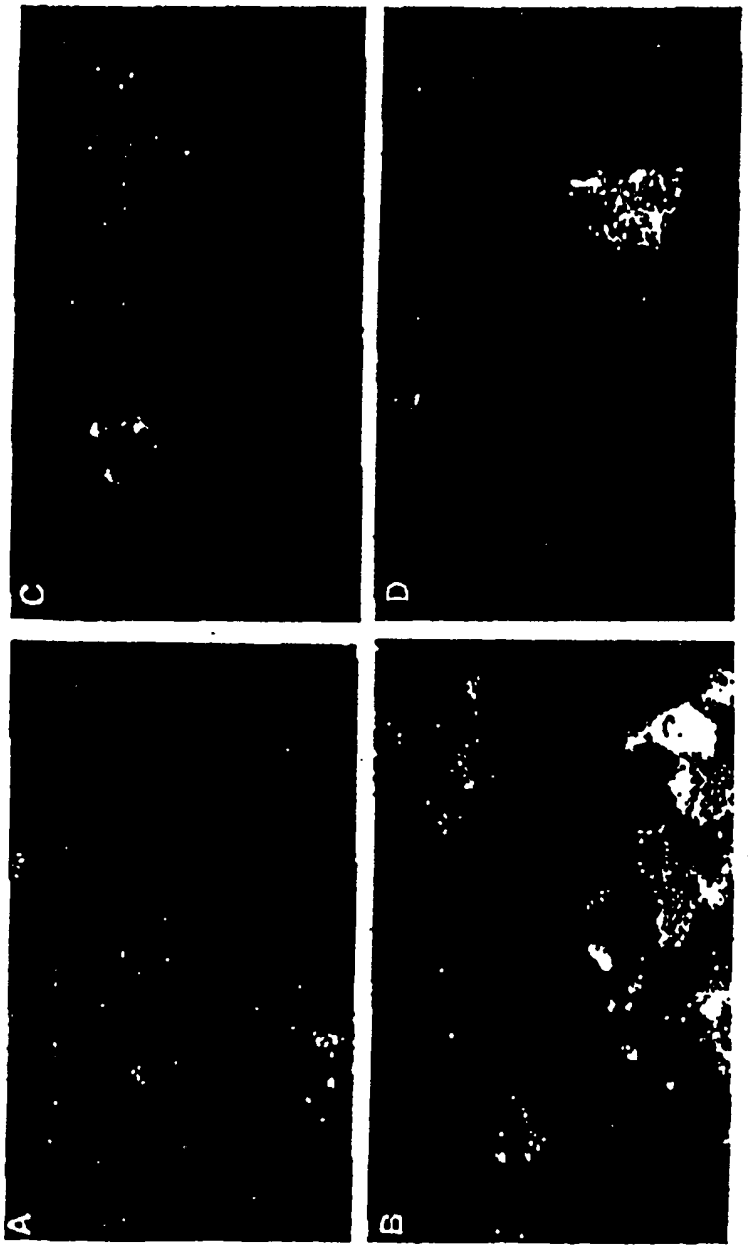


Figure 38C

Figure 38D

Figure 38A

Figure 38B

CLINICAL SIGNS MONKEY C — AGE 7 YEARS

DATE	EXAMINATION	HEART RATE (beats/min)	RESP RATE (breath/min)	TEMPERATURE (Celsius)	WEIGHT (Kg)
5/11/93	NORMAL	112	16	37.8	6.4
5/11/93		INFECTION			
5/14/93	NORMAL	98	14	38.1	
5/18/93	NORMAL	104	16	38.3	
6/4/93	NORMAL	108	16	38.2	
6/18/93	NORMAL	112	16	38.4	
6/24/93	NORMAL	116	18	38.8	
6/24/93		INFECTION			
16/28/93	NORMAL	104	18	37.9	
7/5/93	granulation	116	16	37.4	
7/12/93	NORMAL	114	20	38.3	
9/17/93	NORMAL	108	16	38.3	7

Figure 39A

CLINICAL SIGNS MONKEY D — AGE 7 YEARS

DATE	EXAMINATION	HEART RATE (beats/min)	RESP RATE (breath/min)	TEMPERATURE (Celsius)	WEIGHT (Kg)
5/11/93	NORMAL	108	18	38.3	6.25
5/11/93		INFECTION			
5/14/93	NORMAL	100	20	38.4	
5/18/93	NORMAL	98	20	38.4	
6/4/93	NORMAL	106	18	37.9	
6/18/93	NORMAL	100	19	38.4	
6/24/93	NORMAL	106	16	37.8	
6/24/93		INFECTION			
16/28/93	NORMAL	104	16	37.4	
7/5/93	NORMAL	102	14	38.8	
7/12/93	granulation	114	16	38	
9/17/93	NORMAL	104	16	38.3	6.4

Figure 39B

CLINICAL SIGNS MONKEY E — AGE 11 YEARS

DATE	EXAMINATION	HEART RATE (beats/min)	RESP RATE (breath/min)	TEMPERATURE (Celsius)	WEIGHT (Kg)
5/11/93	NORMAL	120	18	28.3	10
5/11/93		INFECTION			
5/14/93	NORMAL	112	20	37.9	
5/18/93	NORMAL	108	22	38.4	
6/4/93	NORMAL	112	20	38.3	
6/18/93	NORMAL	106	20	38.3	
6/24/93	NORMAL	108	18	38.9	
6/24/93		INFECTION			
16/28/93	NORMAL	112	20	38	
7/5/93	NORMAL	106	22	38.3	
7/12/93	NORMAL	114	16	38	
9/17/93	NORMAL	114	16	38.3	8.75

Figure 39C

Clinical Lab Results From Monkey C

DATE	11-May		11-May	14-May	18-May	4-Jun	18-Jun	24-Jun		24-Jun	12-Jul	17-Sep
WBC/mm3	6.7		9	8.9	7.1	7.9	7.3				10.6	8.1
NEUT/mm3	1850		3990	3060	1480	3550	3450				2210	3950
LYMP/mm3	4460		4220	4770	4780	3640	2670				7270	3770
MONO/mm3	120		520	600	360	420	550				480	340
EOS/mm3	30		110	190	120	80	400				250	70
HEMOG. gr/dl	12.2		12	12.6	12.8	14	13.5				13.7	13.9
HEMATOCR.%	38	F	38	42	41	45	39	S			46	43
PLAT k/mm3	311	I	319	343	338	308	281	E			324	432
ESR	<1	R	1	1	1	0	<1	C			<1	<1
		S						O				
NA mEq/l	149	T	148	147		151	147	N			149	153
K mEq/l	3.6		3.6	2.6		3.6	3.1	D			3.4	3.6
Cl mEq/l	111		106	107		112	108				109	113
CO2 mEq/l	19	I	20	20		22	21	I			19	19
BUN mg/dl	11	N	18	11		14	13	N			16	23
CREAT mg/dl	1.1	F	1	1.2		1.1	1	F			1.1	1.2
GLUCOSE mg/dl	68	E	56	81		67	87	E			74	58
ALB gr/dl	4.7	C	4.3	4.7		4.9	4.2	C			4.5	4.5
T. PROT, gr/dl	7.3	T	6.7	7.1		7.4	6.9	T			7.1	7.4
CALCIUM mg/dl	10	I	9.3	9.9		10.2	9	I			10.1	9.5
PO4 mg/dl	3.3	O	5.9	5.7		2.9	5	O			3.7	3.4
ALK. PH IU/l	117	N	376	375		117	76	N			116	164
TOT BIL mg/dl	0.3		0.2	0.2		0.2	0.1				0.2	0.3
AST IU/l	38		37	45		28	25				45	34
LDH IU/l	601		599	740		277	408				458	220
URIC Ac mg/dl	0.1		0.1	<0.1		0.1	0.1				<0.1	0.1

Figure 40A

Monkey D

Clinical Lab Results From Monkey D										
DATE	11-May	11-May	14-May	18-May	4-Jun	18-Jun	24-Jun	24-Jun	12-Jul	17-Sep
WBC/mm3	7		4.2	9.9	6.7	9.1	6.9		9.4	8.3
NBUT/mm3	2880		1980	3060	1090	6230	1740			3180
LYMP/mm3	3660		4180	6100	4770	1820	4750			3230
MONO/mm3	160		410	340	500	500	190			670
EOS/mm3	50		150	210	110	240	130			210
HEMOG. gr/dl	10.9		13.7	14.7	13.6	13.9	13.8			14.5
HEMATOCR.%	35	F	42	49	44	43	43	S	44	47
PLAT k/mm3	268	I	277	413	369	265	300	E	284	348
ESR	1	R	2	<1	1	0	<1	C	<1	<1
		S						O		
NA mEq/l	147	T	150	150		149	147	N	148	148
K mEq/l	3.5		3.5	3.6		3.5	3.4	D	3.5	3
Cl mEq/l	109		106	110		111	108		109	109
CO2 mEq/l	19	I	20	20		23	20	I	19	16
BUN mg/dl	19	N	18	20		10	16	N	18	12
CREAT mg/dl	1.1	F	1	1.1		1.1	1	F	1	1
GLUCOSEmg/dl	85	E	81	72		92	78	E	66	88
ALB gr/dl	4.3	C	4.7	5.2		4.2	4.6	C	4.5	4.7
T. PROT, gr/dl	6.6	T	7.4	7.8		6.8	8.8	T	7.1	7.6
CALCIUMmg/dl	9.3	I	10.1	10.4		9.6	9	I	10.3	9.5
PO4 mg/dl	6.2	O	3.5	3.6		2.8	5	O	5.6	4.7
ALK. PH IU/l	426	N	104	116		82	337	N	328	101
TOT BIL mg/dl	0.1		0.3	0.2		0.2	0.1		0.1	0.2
AST IU/l	29		32	103		55	27		25	21
LDH IU/l	520		496	912		768	615		252	227
URIC Ac mg/dl	0.1		<0.1	<0.1		0.1	0.1		<0.1	0.1

Figure 40B

Clinical Lab Results From Monkey E

DATE	11-May	11-May	14-May	18-May	4-Jun	18-Jun	24-Jun	24-Jun	12-Jul	17-Sep
WBC/mm3	8.7		7.1		5.3	8.8	8.6		6.9	8.1
NEUT/mm3	4850		2060		3210	4480	2040			2592
LYMP/mm3	3060		4220		1510	3360	5610			5265
MONO/mm3	120		520		280	350	460			162
EOS/mm3	30		110		150	80	170			81
HEMOG. gr/dl	12.9		13.5		13.7	12.6	12.4		13.8	13.9
HEMATOCR.%	40	F	44		42	41	38	S	44	43
PLAT k/mm3	291	I	277		287	291	300	E	269	432
ESR	1	R	1		1	0	<1	C	<1	<1
		S						O		
NA mEq/l	148	T	151	147		148	149	N	148	150
K mEq/l	3		3.3	2.6		3.7	3.6	D	3.1	3.8
Cl mEq/l	110		110	107		110	111		109	110
CO2 mEq/l	16	I	25	20		22	23	I	21	20
BUN mg/dl	8	N	8	11		15	13	N	14	17
CREAT mg/dl	1.1	F	1.2	1.2		1.1	1	F	1	1.2
GLUCOSEmg/dl	115	E	83	102		86	65	E	87	69
ALB gr/dl	4	C	4.2	4.4		4.5	4.8	C	4	4.5
T. PROT, gr/dl	6.7	T	7	7.1		7	7.3	T	6.8	7
CALCIUMmg/dl	9.3	I	9.7	9.4		9.8	9.7	I	9.7	9.4
PO4 mg/dl	3.5	O	4.4	4.2		5.1	3.3	O	4.6	4.1
ALK. PH IU/l	68	N	84	90		393	116	N	75	355
TOT BIL mg/dl	0.2		0.2	0.3		0.1	0.2		0.2	2
AST IU/l	32		29	47		27	28		28	24
LDH IU/l	416		367	571		277	481		247	200
URIC Ac mg/dl	0.1		<0.1	<0.1		0.1	0.1		<0.1	<0.1

Figure 40C

CYTOLOGY MONKEY C

DATE	5/11/93	5/11/93	5/18/93	6/4/93	6/18/93	6/24/93	6/24/93	6/28/93	9/17/93
LEFT NOSTRIL									
Sq. Epith.	68	F	78	63	72	74	S	B	89
Resp. Epith.	30	I	18	84	24	25	E	I	30
Neutrophils	1	R	2	3	2	0	C	O	0
Lymphocytes	1	S	2	0	1	1	O	P	0
Eosinophils	0	T	0	0	1	0	N	S	1
							D	Y	

CYTOLOGY MONKEY D

DATE	5/11/93	5/11/93	5/18/93	6/4/93	6/18/93	6/24/93	6/24/93	7/5/93	9/17/93
LEFT NOSTRIL									
Sq. Epith.	60	F	60	72	72	84	S	B	73
Resp. Epith.	39	I	39	26	25	14	E	I	25
Neutrophils	1	R	1	0	1	2	C	O	2
Lymphocytes	0	S	2	2	1	0	O	P	0
Eosinophils	0	T	0	0	1	0	N	S	0
							D	Y	

CYTOLOGY MONKEY E

DATE	5/11/93	5/11/93	5/18/93	6/4/93	6/18/93	6/24/93	6/24/93	7/12/93	9/17/93
LEFT NOSTRIL									
Sq. Epith.	60	F	60	72	72	84	S	B	73
Resp. Epith.	39	I	39	26	25	14	E	I	25
Neutrophils	1	R	1	0	1	2	C	O	2
Lymphocytes	0	S	2	2	1	0	O	P	0
Eosinophils	0	T	0	0	1	0	N	S	0
							D	Y	

Figure 41

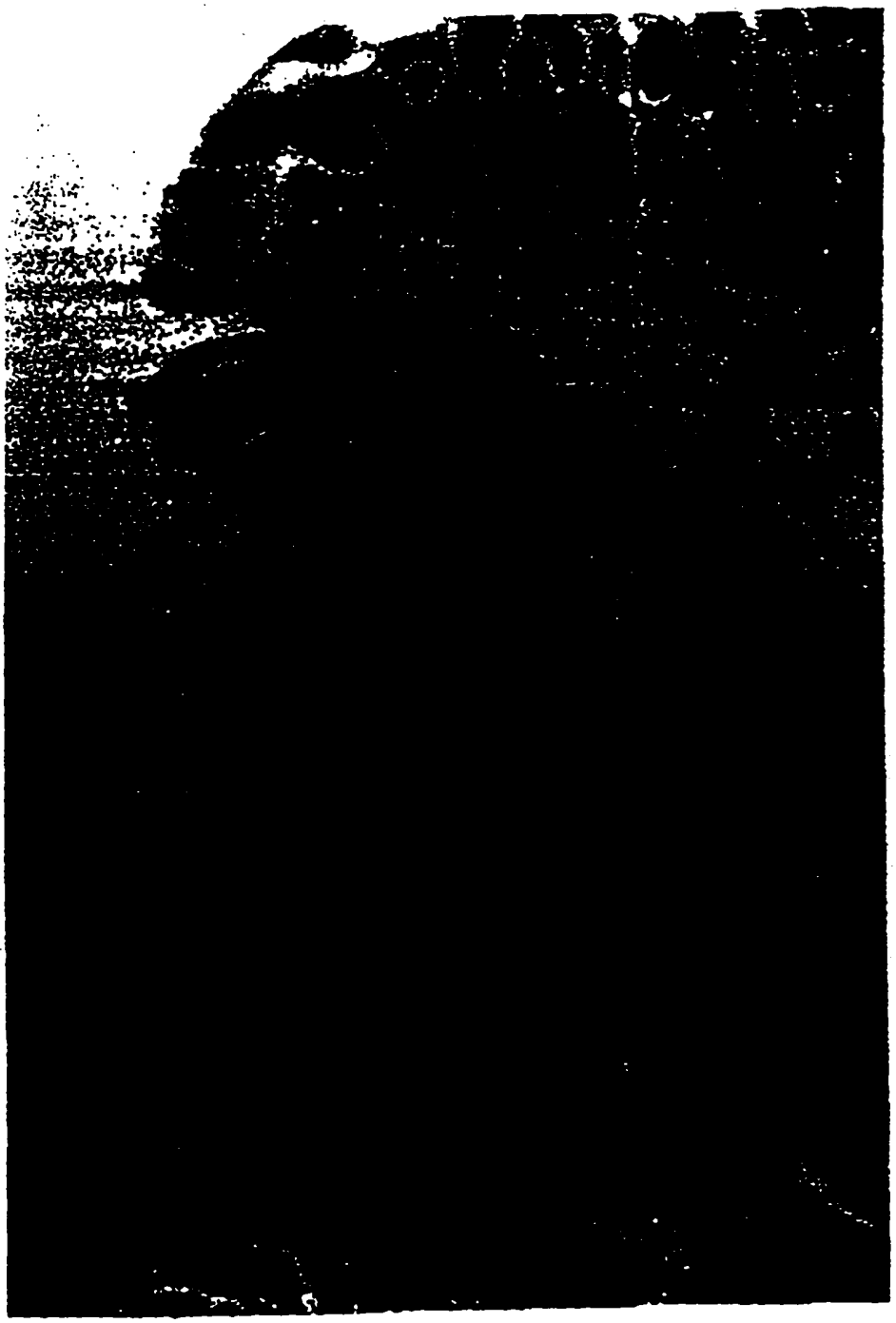


Figure 42

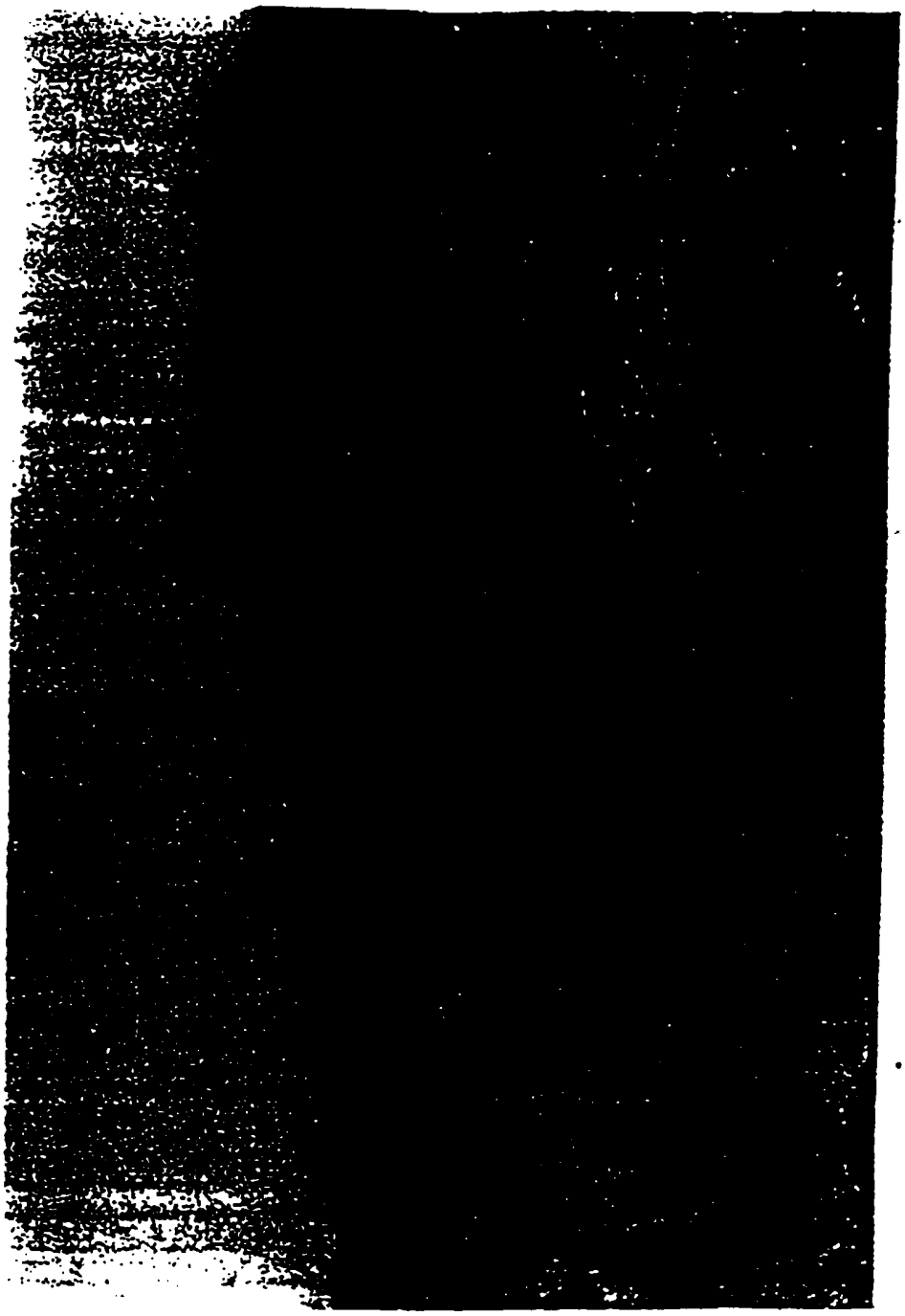


Figure 43

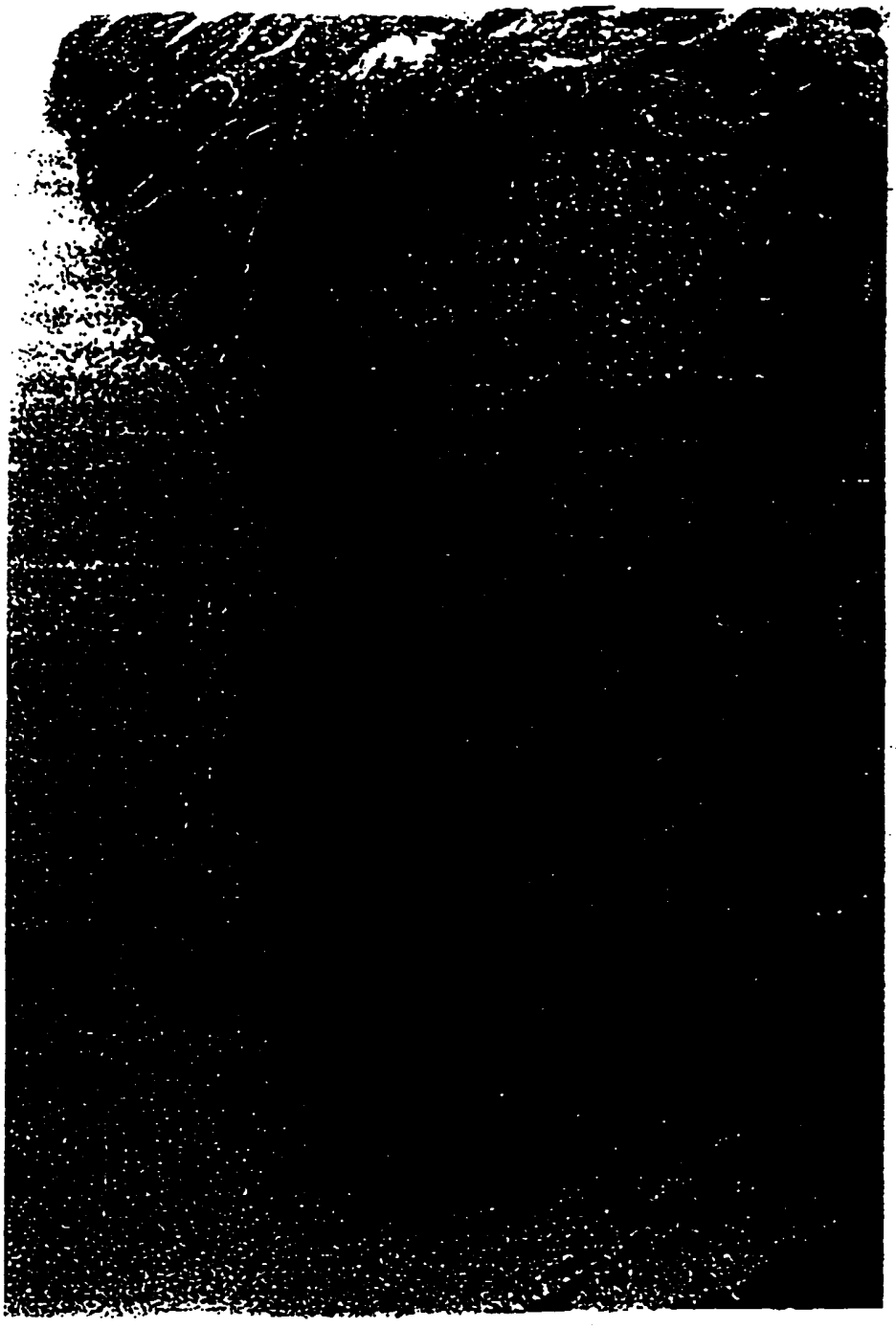


Figure 44

NEUTRALIZING ANTIBODIES o